

UNIVERSIDADE FEDERAL DO PARANÁ

ARIEL ROGELIO VELAZCO CÁRDENAS

PROPRIEDADES DE CONVERGÊNCIA DE UM MÉTODO PQS
ESTABILIZADO PARA PROBLEMAS MATEMÁTICOS COM
CONDIÇÕES DE EQUILÍBRIO

CURITIBA
2019

ARIEL ROGELIO VELAZCO CÁRDENAS

PROPRIEDADES DE CONVERGÊNCIA DE UM MÉTODO PQS ESTABILIZADO
PARA PROBLEMAS MATEMÁTICOS COM CONDIÇÕES DE EQUILÍBRIO

Tese apresentada como requisito parcial à obtenção do grau de Doutor em Matemática, no Curso de Pós-Graduação em Matemática, Setor de Ciências Exatas, da Universidade Federal do Paraná.

Orientador: Prof. Dr. Ademir Alves Ribeiro.
Coorientador: Prof. Dr. José Alberto Ramos Flor.

CURITIBA
2019

C266p

Cárdenas, Ariel Rogelio Velazco

Propriedades de convergência de um método PQS estabilizado para problemas matemáticos com condições de equilíbrio [recurso eletrônico] / Ariel Rogelio Velazco Cárdenas, 2019.

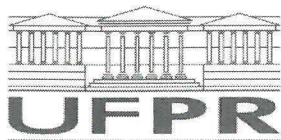
Tese (doutorado) - Programa de Pós-Graduação em Matemática do Setor de Ciências Exatas da Universidade Federal do Paraná

Orientador: Prof. Dr. Ademir Alves Ribeiro

Coorientador: Prof. Dr. José Alberto Ramos Flor

1. Programação quadrática. 2. Análise numérica. I. Universidade Federal do Paraná. II. Ribeiro, Ademir Alves. III. Flor, José Alberto Ramos. IV. Título.

CDD 518.2



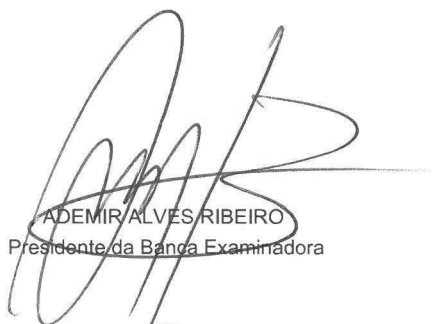
MINISTÉRIO DA EDUCAÇÃO
SETOR CIÊNCIAS EXATAS
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO MATEMÁTICA -
40001016041P1

TERMO DE APROVAÇÃO

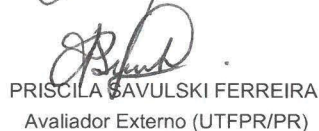
Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em MATEMÁTICA da Universidade Federal do Paraná foram convocados para realizar a arguição da tese de Doutorado de **ARIEL ROGELIO VELAZCO CÁRDENAS** intitulada: **Propriedades de convergência de um método PQS estabilizado para problemas matemáticos com condições de equilíbrio**, após terem inquirido o aluno e realizado a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de doutor está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

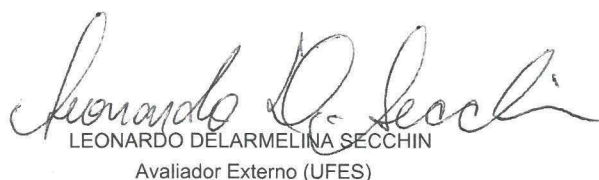
Curitiba, 29 de Janeiro de 2019.



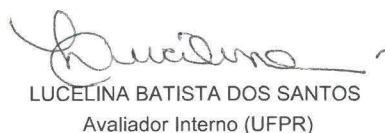
ADEMIR ALVES RIBEIRO
Presidente da Banca Examinadora



PRISCILA SAVULSKI FERREIRA
Avaliador Externo (UTFPR/PR)



LEONARDO DELARMELENA SECCHIN
Avaliador Externo (UFES)



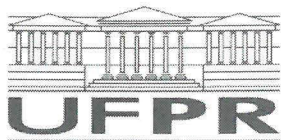
LUCELINA BATISTA DOS SANTOS
Avaliador Interno (UFPR)



LUCAS GARCIA PEDROSO
Avaliador Interno (UFPR)



ROBERTO ANDREANI
Avaliador Externo (UNICAMP)



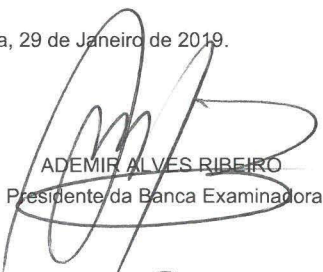
MINISTÉRIO DA EDUCAÇÃO
SETOR CIÊNCIAS EXATAS
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO MATEMÁTICA -
40001016041P1

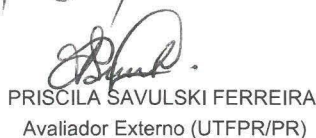
ATA Nº028

ATA DE SESSÃO PÚBLICA DE DEFESA DE DOUTORADO PARA A OBTENÇÃO DO GRAU DE DOUTOR EM MATEMÁTICA

No dia vinte e nove de janeiro de dois mil e dezenove às 14:00 horas, na sala de Reuniões do Setor de Ciências Exatas, Rua Cel. Francisco H. dos Santos, 100 - Jardim das Américas, foram instalados os trabalhos de arguição do doutorando **ARIEL ROGELIO VELAZCO CÁRDENAS** para a Defesa Pública de sua tese intitulada **Propriedades de convergência de um método PQS estabilizado para problemas matemáticos com condições de equilíbrio**. A Banca Examinadora, designada pelo Colegiado do Programa de Pós-Graduação em MATEMÁTICA da Universidade Federal do Paraná, foi constituída pelos seguintes Membros: ADEMIR ALVES RIBEIRO (UFPR), LUCELINA BATISTA DOS SANTOS (UFPR), PRISCILA SAVULSKI FERREIRA (UTFPR/PR), LUCAS GARCIA PEDROSO (UFPR), LEONARDO DELARMELINA SECCHIN (UFES), ROBERTO ANDREANI (UNICAMP). Dando início à sessão, a presidência passou a palavra ao discente, para que o mesmo expusesse seu trabalho aos presentes. Em seguida, a presidência passou a palavra a cada um dos Examinadores, para suas respectivas arguições. O aluno respondeu a cada um dos arguidores. A presidência retomou a palavra para suas considerações finais. A Banca Examinadora, então, reuniu-se e, após a discussão de suas avaliações, decidiu-se pela APROVAÇÃO do aluno. O doutorando foi convidado a ingressar novamente na sala, bem como os demais assistentes, após o que a presidência fez a leitura do Parecer da Banca Examinadora. A aprovação no rito de defesa deverá ser homologada pelo Colegiado do programa, mediante o atendimento de todas as indicações e correções solicitadas pela banca dentro dos prazos regimentais do programa. A outorga do título de doutor está condicionada ao atendimento de todos os requisitos e prazos determinados no regimento do Programa de Pós-Graduação. Nada mais havendo a tratar a presidência deu por encerrada a sessão, da qual eu, ADEMIR ALVES RIBEIRO, lavrei a presente ata, que vai assinada por mim e pelos membros da Comissão Examinadora.

Curitiba, 29 de Janeiro de 2019.


ADEMIR ALVES RIBEIRO
Presidente da Banca Examinadora


PRISCILA SAVULSKI FERREIRA
Avaliador Externo (UTFPR/PR)


LEONARDO DELARMELINA SECCHIN
Avaliador Externo (UFES)


LUCELINA BATISTA DOS SANTOS
Avaliador Interno (UFPR)


LUCAS GARCIA PEDROSO
Avaliador Interno (UFPR)


ROBERTO ANDREANI
Avaliador Externo (UNICAMP)

AGRADECIMENTOS

A minha esposa Mirsa, e minhas filhas Vanessa e Fiorella, pelo amor, carinho, mimos, abraços e suas palavras que tanto me apoiaram e me incentivaram.

Aos meus pais Isidro e Nery, pelo apoio incondicional nos momentos difíceis.

Ao meu orientador, Ademir, pelos muitos momentos de discussões, apoio, tempo, esforço e dedicação empregados, sem os quais a conclusão deste trabalho não teria sido possível.

Ao meu co-orientador, Alberto, pelas ideias e discussões, e pela preocupação por mim.

Aos membros da banca e convidados da pré-defesa, professores Leonardo, Lucelina, Lucas, Roberto, Rodrigo e Priscila obrigado por aceitarem o convite e pelas contribuições sugeridas.

Aos meus amigos que conheci no Brasil ao longo deste período. Em especial, Aura, Elvis, Hector e Leidy, por tantos momentos compartilhados, amizade e companheirismo.

E Finalmente, à CAPES pelo apoio financeiro.

RESUMO

Problemas de Otimização com Condições de Equilíbrio (MPEC) possuem a particularidade de não satisfazer as condições de qualificação usuais para problemas de otimização não linear. Isto representa uma dificuldade ao tentar resolver problemas MPEC com estes métodos. Recentemente, considerando a condição de qualificação MPEC-LICQ, uma adaptação para problemas MPEC da condição de qualificação de independência linear (LICQ) usual, Izmailov, Solodov e Uskov provaram que métodos baseados em Lagrangiano Aumentado de primeira ordem, convergem a pontos C-estacionários, que são mais fracos que pontos KKT. Posteriormente Andreani, Secchin e Silva melhoraram o resultado, mostrando que quando se consideram métodos baseados em Lagrangiano Aumentado de segunda ordem pode-se garantir convergência a pontos pelo menos M-estacionários, condição mais forte que C-estacionariedade, porém também mais fraca que KKT. Além disso mostraram também que considerando a condição MPEC-RCPLD, mais fraca que MPEC-LICQ, e que certa sequência dos multiplicadores é limitada tem-se convergência a pontos S-estacionários, que são equivalentes a KKT. Neste trabalho mostramos que estes resultados não são exclusivos do método de Lagrangiano Aumentado. Apresentamos um método baseado em PQS estabilizada de segunda ordem que pode ser aplicado a problemas MPEC, obtendo resultados equivalentes. Assim, mostramos que quando se considera MPEC-LICQ, o método PQS estabilizado também garante convergência a pontos M-estacionários. Além disso mostramos que considerando MPEC-RCPLD e uma propriedade de limitação vinculada ao multiplicador da restrição de complementaridade também temos convergência a pontos S-estacionários. Testes numéricos são feitos para validar os resultados teóricos.

Palavras-chave: Problemas de Otimização com Condições de Equilíbrio. Programação não linear. Otimização com restrições. Programação Quadrática Sequencial Estabilizada. M-estacionariedade.

ABSTRACT

Mathematical Programs with Equilibrium Constraints (MPEC) have the particularity do not satisfy the usual constraint qualifications for standard nonlinear optimization. This represents a difficulty in attempting to solve MPEC problems with the usual nonlinear optimization methods. Recently, considering the MPEC-LICQ constraint qualification, an adaptation to MPEC problems of the usual linear independence constraint qualification (LICQ), Izmailov, Solodov and Uskov proved that first order augmented Lagrangian methods converge to C-stationary points, which are weaker than KKT points. Later on Andreani, Secchin and Silva improved this result, showing that when considering second order augmented Lagrangian methods, it can be guaranteed convergence to at least M-stationary points, which are stronger than C-stationary, but still weaker than KKT. In addition, they also showed that considering the MPEC-RCPLD constraint qualification, which is weaker than MPEC-LICQ, and assuming that a certain multiplier sequence is bounded, it can be proved the convergence to S-stationary points, which are equivalent to KKT. In this work we show that these results are not exclusive of the Augmented Lagrangian method. We provide a method based on second order stabilized Sequential Quadratic Programming, which can be applied to MPEC problems, achieving equivalent results. Thus we show that when MPEC-LICQ is considered, the stabilized SQP method also guarantees convergence to M-stationary points. Moreover, we show that considering MPEC-RCPLD and a boundedness property related to the complementarity constraint multiplier, the method also has convergence to S-stationary points. Numerical tests were performed to validate the theoretical results.

Keywords: Mathematical Programs with Equilibrium Constraints. Non-linear programming. Constrained optimization. stabilized Sequential Quadratic Programming. M-stationarity.

LISTA DE SÍMBOLOS

$\mathcal{T}(x^*)$: Cone tangente.

$\mathcal{T}^{lin}(x^*)$: Cone viável linearizado.

C^* : Dual do conjunto C .

I_{0+} : Conjunto de índices $\{i \mid G_i(x^*) = 0, H_i(x^*) > 0\}$ do Problema (1).

I_{+0} : Conjunto de índices $\{i \mid G_i(x^*) > 0, H_i(x^*) = 0\}$ do Problema (1).

I_{00} : Conjunto de índices $\{i \mid G_i(x^*) = 0, H_i(x^*) = 0\}$ do Problema (1).

$NLP_*(\beta_1, \beta_2)$: Problema não linear (β_1, β_2) -restrito.

$\mathcal{T}_{NLP_*(\beta_1, \beta_2)}(x^*)$: Cone tangente para problema $NLP_*(\beta_1, \beta_2)$.

$\mathcal{T}_{NLP_*(\beta_1, \beta_2)}^{lin}(x^*)$: Cone viável linearizado problema $NLP_*(\beta_1, \beta_2)$.

$\mathcal{T}_{MPEC}^{lin}(x^*)$: Cone viável linearizado MPEC.

λ_i^G, λ_i^H : Multiplicadores MPEC associados às restrições G_i e H_i do Problema (1).

$J(x)$: Matriz Jacobiana das restrições $c(x)$ do Problema (3.1).

$H(x, y)$: Hessiana da Lagrangiana do Problema (3.1).

$r(x, y)$: Função proximal para o Problema (3.1).

y_k^E : Multiplicador estimado do Algoritmo 5.

μ_k^R : Parâmetro de regularização do Algoritmo 5.

μ_k : Parâmetro de penalidade do Algoritmo 5.

$M(x, y; y^E, \mu)$: Função Lagrangiana aumentada primal-dual.

$Q_k(v, y_k^E, \mu)$: Modelo quadrático local de $M(x, y; y^E, \mu)$.

$\mathcal{A}(x)$: Conjunto ativos para o Problema (3.1).

$\mathcal{F}(x)$: Conjunto primal-dual livre para o Problema (3.1).

$A_\epsilon(x)$: Conjunto ϵ -ativo para o Problema (3.1).

$F_\epsilon(x)$: Conjunto ϵ -livre para o Problema (3.1).

ϕ_O, ϕ_V : Funções de mérito associadas às iterações de tipo O e V.

\mathcal{S} : Conjunto de índices das iterações de tipo O e V.

\hat{y}_k : Multiplicador aproximado associado à restrição g do Problema (4.6).

\hat{y}_k^G, \hat{y}_k^H e \hat{y}_k^0 : Multiplicadores associados às restrições do Problema (4.1).

z_k^G, z_k^H e z_k^0 : Variáveis de folga associadas às restrições do Problema (4.1).

$y_k^{E,G}, y_k^{E,H}$ e $y_k^{E,0}$: Multiplicadores estimados associados às restrições do Problema (4.1).

γ^k, ν^k : Multiplicadores MPEC associados às restrições G_i e H_i do Problema (4.1).

I_0 : Conjunto I_{00} do Problema (4.1).

I_+ : Conjunto I_{0+} do Problema (4.1).

LISTA DE SIGLAS E ABREVIACÕES

MPEC : Problemas de otimização com condições de equilíbrio.

CQ : Condição de qualificação.

KKT : Karush Kun Tucker.

LICQ : Condição de qualificação de independência linear.

MFCQ : Condição de qualificação de Mangasarian-Fromovitz.

ACQ : Condição de qualificação de Abadie.

GCQ : Condição de qualificação de Guignard.

SCQ : Condição de qualificação de Slater.

SOSC : Condições suficientes de segunda ordem.

RCPLD : Dependência linear positiva constante relaxada.

PQS : Programação Quadrática Sequencial.

PQSe : Programação Quadrática Sequencial estabilizada.

AKKT : Aproximadamente KKT.

OCE : Ordem de convergência estimada.

Sumário

Introdução	3
1 Conceitos básicos de programação não linear	5
1.1 Condições de qualificação	6
1.2 Condições de qualificação usuais aplicadas a problemas MPEC	9
2 Condições de qualificação e estacionariedade em MPEC	18
2.1 Condições de qualificação MPEC usuais	18
2.1.1 MPEC-LICQ e MPEC-MFCQ	18
2.1.2 Cone Viável Linearizado MPEC	19
2.1.3 MPEC-ACQ e MPEC-GCQ	21
2.1.4 MPEC-ACQ e MPEC-MFCQ	22
2.2 Relação entre as CQs MPEC e as CQs usuais	26
2.2.1 MPEC-LICQ e GCQ	27
2.2.2 MPEC-ACQ e GCQ	29
2.3 Estacionariedade em MPEC	34
2.4 Condições de qualificação para MPEC e estacionariedade	37
2.4.1 MPEC-LICQ e estacionariedade	37
2.4.2 MPEC-MFCQ e estacionariedade	39
2.4.3 MPEC-GCQ e estacionariedade	40
3 Método PQS estabilizado	48
3.1 Conceitos preliminares	49
3.2 Fundamentação do método	51
3.3 Descrição do Algoritmo	54
3.3.1 Estimando uma direção de curvatura negativa de Q_k	55
3.3.2 Atualizando o multiplicador estimado y_k^E e o parâmetro de regularização μ_k^R	56
3.3.3 Definição da direção de busca linear	58
3.3.4 Busca linear flexível	61
3.3.5 Critério de parada	63
3.4 Propriedades do Algoritmo 5	65

4	Convergência para pontos S-estacionários e M-estacionários	72
4.1	Pontos Aproximadamente KKT	73
4.2	Fórmula explícita para os multiplicadores de Lagrange	78
4.3	Propriedades de convergência do algoritmo para problemas com desigualdade	81
4.4	Resultados de convergência para MPECs	83
4.4.1	Convergência para pontos S-estacionários	85
4.4.2	Convergência para pontos M-estacionários	90
4.5	Testes numéricos	102
	Conclusões	108
	Referências	116
	Algoritmos	117

Introdução

Estamos interessados em problemas do tipo

$$\begin{aligned} \min \quad & f(x) \\ \text{s.a.} \quad & g(x) \leq 0, \quad h(x) = 0 \\ & G(x) \geq 0, \quad H(x) \geq 0, \quad G(x)^T H(x) = 0 \end{aligned} \tag{1}$$

onde $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $h : \mathbb{R}^n \rightarrow \mathbb{R}^p$, $G : \mathbb{R}^n \rightarrow \mathbb{R}^l$ e $H : \mathbb{R}^n \rightarrow \mathbb{R}^l$ são funções continuamente diferenciáveis. Estes problemas recebem o nome de “*problemas de otimização com condições de equilíbrio*” (MPECs) ou “*problemas de otimização com condições de complementaridade*” (MPCCs), sendo ambos usuais na literatura. A nomenclatura “MPEC” se deve ao fato de que originalmente aparecerem em problemas de equilíbrio de economia e engenharia, e será usada no presente trabalho por ter uma pronúncia mais chamativa. Na atualidade MPECs possuem aplicações variadas em economia, jogos, equilíbrio químico, desenho de estruturas, fricção, análise de risco, controle urbano, etc; ver por exemplo [10, 13, 16, 17, 19, 33, 34, 53, 55, 68, 70] e suas referências.

No entanto, este tipo de problema é geralmente difícil de lidar porque suas restrições não cumprem algumas das condições de qualificação usuais, como por exemplo a de Mangasarian-Fromovitz (MFCQ) [75]. Isto significa que não há garantia de que os métodos tradicionais de otimização aplicados ao MPEC convirjam a pontos KKT, onde a principal dificuldade reside nos sinais dos multiplicadores, não necessariamente positivos, associados às restrições H_i e G_i quando as duas são ativas ao mesmo tempo. Muita pesquisa tem sido feita durante as duas últimas décadas acerca das condições de otimalidade para MPECs, introduzindo-se novos conceitos de estacionariedade como C-estacionariedade (Clarke), M-estacionariedade (Mordukhovich) e S-estacionariedade (forte), entre outros; ver por exemplo [21, 34, 35, 46, 53, 56, 57, 58, 65, 74, 75, 76, 77] e suas referências. Entre estes conceitos o de M-estacionariedade é mais forte que C-estacionariedade e mais fraco que S-estacionariedade, que é equivalente a KKT [23]. Porém, as restrições do problema não linear devem satisfazer uma condição de qualificação para que um minimizador satisfaça algum conceito de estacionariedade. Nesse sentido, condições de qualificação específicas para MPEC, como MPEC-LICQ, MPEC-MFCQ, MPEC-ACQ e MPEC-GCQ, foram introduzidas; ver por exemplo [34, 37, 38, 46, 61, 65, 67, 74] e suas referências.

Paralelamente, algoritmos específicos para resolver MPEC foram propostos usando uma série de abordagens baseadas em programação quadrática sequencial, penalidade, relaxação e conjunto ativo; ver por exemplo, [26, 27, 39, 40, 48, 51, 53, 55, 71, 72] e suas referências. Além disso, muitos métodos específicos para MPEC foram desenvolvidos, ver [18, 44, 45, 47, 50, 52, 66], porém apenas a convergência a pontos C-estacionários pode ser garantida, a menos de hipóteses fortes. Como consequência, o uso de métodos convencionais é considerado. Izmailov, Solodov e Uskov [40] provaram que métodos baseados em Lagrangiano Aumentado de primeira ordem, considerando MPEC-LICQ, convergem para pontos C-estacionários. Posteriormente, Andreani, Secchin e Silva [1] melhoraram o resultado mostrando que quando se consideram métodos baseados em Lagrangiano Aumentado de segunda ordem, supondo MPEC-LICQ, pode-se garantir convergência para pontos que são pelo menos M-estacionários. Além disso em [1] também se mostra que considerando MPEC-RCPLD e uma propriedade de limitação vinculada ao multiplicador da restrição de complementaridade, tem-se convergência para pontos S-estacionários.

Existem muitos outros métodos de segunda ordem para otimização não linear geral, dentre os quais o método desenvolvido por Gill, Kungurtsev e Robinson [30] é baseado em programação quadrática sequencial estabilizada. Eles apresentam o algoritmo PDSQP2 que possui convergência superlinear, mostrando também que em seus testes numéricos aplicados a problemas MPEC o algoritmo resolveu 78 de 86 problemas obtidos de uma coleção proposta por Sven Leyffer (Argonne National Laboratory).

Motivados por estes resultados, nosso interesse é analisar o comportamento do método de programação quadrática sequencial estabilizada de segunda ordem quando aplicado a problemas MPEC. Para isso fazemos uma modificação ao algoritmo PDSQP2 e analisamos as propriedades deste novo algoritmo quando aplicado a problemas MPEC. Mostramos que assumindo MPEC-LICQ, este novo algoritmo tem convergência para pontos M-estacionários. Além disso, similarmente ao feito em [1], mostramos que considerando MPEC-RCPLD e uma propriedade de limitação vinculada ao multiplicador da restrição de complementaridade, tem-se convergência para pontos S-estacionários. A seguir falaremos sobre a estrutura do texto.

Primeiramente, no Capítulo 1 fazemos uma revisão de conceitos de programação não linear, com ênfase nas condições de qualificação usuais e seu comportamento em problemas MPEC. No Capítulo 2 fornecemos os conceitos associados aos problemas MPEC e suas relações com os conceitos de programação não linear usual. No Capítulo 3 propomos um método PQS e discutimos suas propriedades. No Capítulo 4 apresentamos os resultados teóricos obtidos relacionados à convergência do algoritmo quando aplicado a problemas MPEC. Testes numéricos são realizados corroborando nossa teoria.

Capítulo 1

Conceitos básicos de programação não linear

Consideremos o problema de otimização não linear

$$\begin{aligned} \min \quad & f(x) \\ \text{s.a} \quad & g(x) \leq 0 \\ & h(x) = 0 \end{aligned} \tag{1.1}$$

onde $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $h : \mathbb{R}^n \rightarrow \mathbb{R}^p$ e $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ são funções continuamente diferenciáveis. A Lagrangiana associada ao Problema (1.1) é dada por

$$\mathcal{L}(x, \lambda^h, \lambda^g) = f(x) + \sum_{i=1}^p \lambda_i^h h_i(x) + \sum_{j=1}^m \lambda_j^g g_j(x)$$

onde as componentes do vetor $\lambda = (\lambda^h, \lambda^g)$ são conhecidas como multiplicadores de Lagrange. A maioria dos métodos desenvolvidos para a resolução do problema anterior buscam encontrar pontos que satisfazem as condições de otimalidade de Karush-Kunh-Tucker (KKT), ou seja procuram encontrar pontos estacionários. Formalmente, se x^* é um ponto viável do Problema (1.1), dizemos que x^* satisfaz as condições de Karush-Kunh-Tucker (KKT) se existem vetores λ^h e λ^g tais que

$$\begin{aligned} \nabla_x \mathcal{L}(x, \lambda^h, \lambda^g) &= \nabla f(x^*) + \sum_{i=1}^p \lambda_i^h \nabla h_i(x^*) + \sum_{j=1}^m \lambda_j^g \nabla g_j(x^*) = 0, \\ \lambda_j^g &\geq 0, \quad j = 1, \dots, m, \\ \lambda_j^g g_j(x^*) &= 0, \quad i = 1, \dots, m. \end{aligned}$$

A seguir faremos uma revisão das condições de qualificação usuais e suas propriedades, relacionando-as ao problema MPEC.

1.1 Condições de qualificação

Como é conhecido na literatura, um minimizador do Problema (1.1) será estacionário se as restrições do problema satisfazem alguma condição de qualificação (CQ). Descrevemos a seguir algumas das principais CQs e a relação que existe entre elas.

Seja Ω o conjunto viável do Problema (1.1) e $\mathcal{I}_g = \{i \mid g_i(x^*) = 0\}$ o conjunto dos índices das restrições de desigualdade ativas.

Definição 1.1. Dizemos que a condição de qualificação de independência linear (LICQ) é satisfeita em $x^* \in \Omega$ quando o conjunto formado pelos gradientes

$$\{\nabla g_i(x^*), \nabla h_j(x^*) \mid i \in \mathcal{I}_g \text{ e } j = 1, \dots, p\} \quad (1.2)$$

é linearmente independente.

Definição 1.2. Dizemos que a condição de qualificação de Mangasarian-Fromovitz (MFCQ) é satisfeita em $x^* \in \Omega$ quando o conjunto formado pelos gradientes

$$\{\nabla h_i(x^*) \mid i = 1, \dots, p\}$$

é linearmente independente e existe um vetor $d \in \mathbb{R}^n$ tal que

$$\begin{aligned} \nabla g_i(x^*)^T d &< 0, \forall i \in \mathcal{I}_g, \\ \nabla h_i(x^*)^T d &= 0, \forall i = 1, \dots, p. \end{aligned} \quad (1.3)$$

Com relação a estas definições temos que a condição MFCQ é mais fraca que LICQ, como podemos ver no seguinte resultado:

Teorema 1.3. [11] Seja x^* um ponto viável do Problema (1.1). Se x^* satisfaz LICQ, então x^* satisfaz MFCQ.

Outras duas condições de qualificação usuais estão relacionadas com alguns cones especiais, definidos a seguir.

Definição 1.4. Para $x^* \in \Omega$, definimos o cone tangente em x^* como o conjunto

$$\mathcal{T}(x^*) = \left\{ d \in \mathbb{R}^n \mid \exists \{x^k\} \subseteq \Omega, \text{ com } x^k \rightarrow x^* \text{ e } \frac{x^k - x^*}{\|x^k - x^*\|} \rightarrow \frac{d}{\|d\|} \right\} \cup \{0\} \quad (1.4)$$

e o cone viável linearizado em x^* como o conjunto

$$\begin{aligned} \mathcal{T}^{lin}(x^*) = \{d \in \mathbb{R}^n \mid \nabla g_i(x^*)^T d \leq 0, \forall i \in \mathcal{I}_g \text{ e} \\ \nabla h_i(x^*)^T d = 0, \forall i = 1, \dots, p\}. \end{aligned} \quad (1.5)$$

Observação 1.5. A definição do cone tangente é um conceito local, isto é, basta somente considerar a sequência $\{x^k\}$ numa vizinhança relativa $B \cap \Omega$ de x^* , de fato

$$\mathcal{T}(x^*) = \mathcal{T}(x^*, B \cap \Omega) = \left\{ d \in \mathbb{R}^n \mid \exists \{x^k\} \text{ em } B \cap \Omega, x^k \rightarrow x^* \text{ e } \frac{x^k - x^*}{\|x^k - x^*\|} \rightarrow \frac{d}{\|d\|} \right\} \cup \{0\}.$$

Além disso, dadas duas vizinhanças relativas $B \cap \Omega$ e $A \cap \Omega$ de x^* , então temos

$$\mathcal{T}(x^*, (A \cup B) \cap \Omega) = \mathcal{T}(x^*, A \cap \Omega) \cup \mathcal{T}(x^*, B \cap \Omega).$$

Definição 1.6. Dado $\mathcal{C} \subseteq \mathbb{R}^n$, definimos o dual do conjunto \mathcal{C} como o conjunto

$$\mathcal{C}^* = \{v \in \mathbb{R}^n \mid v^T d \geq 0, \forall d \in \mathcal{C}\}. \quad (1.6)$$

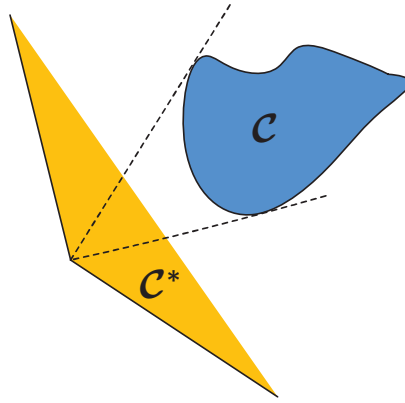


Figura 1.1: Cone dual \mathcal{C}^* do conjunto \mathcal{C} .

Destacamos que na literatura existe um conceito similar chamado cone polar $P(\mathcal{C})$ o qual está relacionado ao cone polar por $P(\mathcal{C}) = -\mathcal{C}^*$. No que se refere ao cone tangente e ao cone viável linearizado destacamos os resultados a seguir:

Lema 1.7. [60, Capítulo 7] Dado $x^* \in \Omega$

- (a) O cone $\mathcal{T}^{lin}(x^*)$ é convexo, fechado e não vazio;
- (b) $\mathcal{T}(x^*) \subseteq \mathcal{T}^{lin}(x^*)$;
- (c) Se x^* é um minimizador local do Problema (1.1), então $\nabla f(x^*)^T d \geq 0$, para todo $d \in \mathcal{T}(x^*)$.

Para o caso do cone dual, destacamos os seguintes resultados que se encontram em [60, Capítulo 7] considerando sua relação com o cone polar.

Lema 1.8. Sejam $\mathcal{C}, \tilde{\mathcal{C}} \subseteq \mathbb{R}^n$

- (a) \mathcal{C}^* é um cone, convexo e fechado;

- (b) $\mathcal{C} \subseteq \mathcal{C}^{**}$;
- (c) (**Farkas**) Se \mathcal{C} é um cone convexo fechado não vazio. Então $\mathcal{C}^{**} = \mathcal{C}$;
- (d) Se \mathcal{C} e $\tilde{\mathcal{C}}$ são cones não vazios, então $(\mathcal{C} \cup \tilde{\mathcal{C}})^* = \mathcal{C}^* \cap \tilde{\mathcal{C}}^*$;
- (e) Se $\mathcal{C} \subseteq \tilde{\mathcal{C}}$ então $\tilde{\mathcal{C}}^* \subseteq \mathcal{C}^*$.

Lema 1.9. [60, Lemas 7.18 e 7.19] Dados $a_1, \dots, a_k, b_1, \dots, b_l \in \mathbb{R}^n$, considere os cones

$$\mathcal{K}_1 = \{d \in \mathbb{R}^n \mid a_i^T d \geq 0 \quad \forall i = 1, \dots, k, \quad b_j^T d = 0 \quad \forall j = 1, \dots, l\}$$

e

$$\mathcal{K}_2 = \{v \in \mathbb{R}^n \mid v = \sum_{i=1}^k \alpha_i a_i + \sum_{j=1}^l \beta_j b_j, \quad \alpha_i \geq 0 \quad \forall i = 1, \dots, k \text{ e } \beta_j \in \mathbb{R}\}.$$

Então $\mathcal{K}_1 = \mathcal{K}_2^*$ e $\mathcal{K}_2 = \mathcal{K}_1^*$.

Os cones anteriores permitem definir as seguintes condições de qualificação usuais na literatura:

Definição 1.10. Dizemos que a condição de qualificação de Abadie (ACQ) é satisfeita em $x^* \in \Omega$ se

$$\mathcal{T}(x^*) = \mathcal{T}^{lin}(x^*),$$

e que condição de qualificação de Guignard (GCQ) é satisfeita em x^* se

$$\mathcal{T}(x^*)^* = \mathcal{T}^{lin}(x^*)^*. \quad (1.7)$$

Obviamente ACQ implica GCQ, e também pode-se demonstrar o seguinte resultado:

Teorema 1.11. [60, Teorema 7.43] Se $x^* \in \Omega$ satisfaz MFCQ, então satisfaz ACQ.

Quando a restrição h é linear e g é convexa temos a seguinte condição de qualificação:

Definição 1.12. A condição de qualificação de Slater (SCQ) é satisfeita em $x^* \in \Omega$ se g é convexa, h linear e existe um vetor $\hat{x} \in \Omega$ tal que

$$g(\hat{x}) < 0 \quad \text{e} \quad h(\hat{x}) = 0. \quad (1.8)$$

SCQ se relaciona com as condições anteriores mediante o seguinte resultado:

Teorema 1.13. [28, Teorema 2.45] Se $x^* \in \Omega$ satisfaz SCQ, então satisfaz ACQ.

As implicações abaixo resumem as relações que existem entre estas CQs:

$$LICQ \implies MFCQ \implies ACQ \implies GCQ \quad (1.9)$$

$$SCQ \implies ACQ.$$

As condições do KKT são de primeira ordem, uma vez que são formuladas em termos de derivadas de primeira ordem das funções envolvidas no problema. No entanto, existem condições de segunda ordem que levam em conta informações sobre a curvatura dessas funções. Em particular, estamos interessados na condição a seguir:

Definição 1.14. *Considerando no Problema (1.1) que $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $h : \mathbb{R}^n \rightarrow \mathbb{R}^p$ e $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ são funções duas vezes diferenciáveis. O vetor $(x^*, \lambda^h, \lambda^g)$ satisfaz as condições suficientes de segunda ordem (SOSC) para o Problema (1.1) se satisfaz as condições de KKT e*

$$p^T \nabla_{xx}^2 \mathcal{L}(x, \lambda^h, \lambda^g) p > 0 \text{ para todo } p \in \mathcal{C}(x^*) \setminus \{0\},$$

onde $\mathcal{C}(x^*)$ é o cone crítico

$$\mathcal{C}(x^*) = \{p \mid \nabla f(x^*)^T p = 0, \nabla h_i(x^*)^T p = 0, \nabla g_j(x^*)^T p \leq 0 \text{ para } i = 1, \dots, p \text{ e } j \in \mathcal{I}_g\}.$$

Teorema 1.15. [41, Teorema 1.20] *Se o vetor $(x^*, \lambda^h, \lambda^g)$ satisfaz SOSC para o Problema (1.1), então x^* é um minimizador local estrito.*

1.2 Condições de qualificação usuais aplicadas a problemas MPEC

Voltando ao Problema (1), a restrição $G(x)^T H(x) = 0$, chamada restrição de complementaridade, é a que faz com que o Problema (1) não cumpra algumas das CQs usuais apresentadas na Seção 1.1. Note que considerando também as restrições $G(x) \geq 0$ e $H(x) \geq 0$ temos que $G_i(x) = 0$ ou $H_i(x) = 0$ para todo $i = 1, \dots, l$. Desta forma, para uma melhor análise consideremos x^* um ponto viável do Problema (1) e os seguintes conjuntos

$$\begin{aligned} I_{0+} &= I_{0+}(x^*) = \{i \mid G_i(x^*) = 0, H_i(x^*) > 0\}, \\ I_{+0} &= I_{+0}(x^*) = \{i \mid G_i(x^*) > 0, H_i(x^*) = 0\}, \\ I_{00} &= I_{00}(x^*) = \{i \mid G_i(x^*) = 0, H_i(x^*) = 0\}, \end{aligned} \quad (1.10)$$

que podem ser melhor visualizados no seguinte quadro:

	$H_i(x^*) = 0$	$H_i(x^*) > 0$
$G_i(x^*) = 0$	$I_{00}(x^*)$	$I_{0+}(x^*)$
$G_i(x^*) > 0$	$I_{+0}(x^*)$	não existe

Como veremos mais adiante na Seção 2.3, os conceitos de estacionariedade em MPEC estão reacionados ao sinal dos multiplicadores no conjunto I_{00} , por esta razão vamos supor que $I_{00} \neq \emptyset$.

Um dos problemas dos MPECs é que os pressupostos que se requerem para garantir a convergência para pontos estacionários da maioria dos métodos usuais não são satisfeitos, mesmo no caso onde todas as restrições são lineares. A seguir veremos o comportamento das CQs usuais vistas na seção anterior quando aplicadas a problemas MPEC.

Proposição 1.16. [75, Proposição 1.1] *Seja x^* um ponto viável do Problema MPEC (1). Então x^* não satisfaz MFCQ, em particular não satisfaz LICQ.*

Demonstração. Consideremos $\theta(x) = G(x)^T H(x)$ e suponhamos por absurdo que x^* satisfaz MFCQ. Então o conjunto

$$\{\nabla\theta(x^*), \nabla h_i(x^*) \mid i = 1, \dots, p\} \quad (1.11)$$

é linearmente independente e existe um vetor $d \in \mathbb{R}^n$ tal que

$$\begin{aligned} \nabla g_i(x^*)^T d &< 0, \quad \forall i \in \mathcal{I}_g \\ \nabla h_i(x^*)^T d &= 0, \quad \forall i = 1, \dots, p \\ \nabla G_i(x^*)^T d &> 0, \quad \forall i \in I_{0+} \cup I_{00} \\ \nabla H_i(x^*)^T d &> 0, \quad \forall i \in I_{+0} \cup I_{00} \\ \nabla\theta(x^*)^T d &= 0. \end{aligned} \quad (1.12)$$

Caso $I_{0+} = I_{+0} = \emptyset$ temos que

$$\nabla\theta(x^*) = \sum_{i=1}^l [G_i(x^*) \nabla H_i(x^*) + H_i(x^*) \nabla G_i(x^*)] = 0,$$

o que contradiz a independência linear do conjunto (1.11). Assumindo agora que I_{0+} ou I_{+0} são diferentes do vazio, temos que

$$\begin{aligned} \nabla\theta(x^*)^T d &= \sum_{i=1}^l [G_i(x^*) \nabla H_i(x^*) + H_i(x^*) \nabla G_i(x^*)]^T d \\ &= \sum_{i \in I_{+0}} \underbrace{G_i(x^*)^T d}_{>0} \underbrace{\nabla H_i(x^*)^T d}_{>0} + \sum_{i \in I_{0+}} \underbrace{H_i(x^*)^T d}_{>0} \underbrace{\nabla G_i(x^*)^T d}_{>0} \neq 0, \end{aligned}$$

o que contradiz a última relação de (1.12). Logo, o ponto x^* não satisfaz MFCQ e portanto também não cumpre LICQ. \square

Proposição 1.17. *Se x^* é um ponto viável do Problema MPEC (1), então x^* não satisfaz SCQ.*

Demonstração. SCQ nunca é satisfeita pois não existe $\hat{x} \in \mathbb{R}^n$ tal que $G(\hat{x}) > 0$, $H(\hat{x}) > 0$, $G(\hat{x})^T H(\hat{x}) = 0$. \square

No que se refere a ACQ, precisamos primeiro observar algumas das propriedades dos cones $\mathcal{T}(x^*)$ e $\mathcal{T}^{lin}(x^*)$ associados aos problemas MPEC. Neste caso o cone viável linearizado $\mathcal{T}^{lin}(x^*)$ tem a forma

$$\begin{aligned} \mathcal{T}^{lin}(x^*) = \{d \in \mathbb{R}^n \mid & \nabla g_i(x^*)^T d \leq 0, \quad \forall i \in \mathcal{I}_g \\ & \nabla h_i(x^*)^T d = 0, \quad \forall i = 1, \dots, p \\ & \nabla G_i(x^*)^T d \geq 0, \quad \forall i \in I_{0+} \cup I_{00} \\ & \nabla H_i(x^*)^T d \geq 0, \quad \forall i \in I_{+0} \cup I_{00} \\ & \nabla \theta(x^*)^T d = 0\}, \end{aligned}$$

onde $\theta(x) = G(x)^T H(x)$. Porém, como

$$\begin{aligned} \nabla \theta(x^*)^T d &= \sum_{i=1}^l [G_i(x^*) \nabla H_i(x^*) + H_i(x^*) \nabla G_i(x^*)]^T d \\ &= \sum_{i \in I_{+0}} \underbrace{G_i(x^*)}_{>0} \underbrace{\nabla H_i(x^*)^T d}_{\geq 0} + \sum_{i \in I_{0+}} \underbrace{H_i(x^*)}_{>0} \underbrace{\nabla G_i(x^*)^T d}_{\geq 0} = 0, \end{aligned}$$

então o cone $\mathcal{T}^{lin}(x^*)$ pode ser expresso como:

$$\begin{aligned} \mathcal{T}^{lin}(x^*) = \{d \in \mathbb{R}^n \mid & \nabla g_i(x^*)^T d \leq 0, \quad \forall i \in \mathcal{I}_g \\ & \nabla h_i(x^*)^T d = 0, \quad \forall i = 1, \dots, p \\ & \nabla G_i(x^*)^T d = 0, \quad \forall i \in I_{0+} \\ & \nabla H_i(x^*)^T d = 0, \quad \forall i \in I_{+0} \\ & \nabla G_i(x^*)^T d \geq 0, \quad \forall i \in I_{00} \\ & \nabla H_i(x^*)^T d \geq 0, \quad \forall i \in I_{00}\}. \end{aligned} \tag{1.13}$$

Também será necessário introduzir o seguinte problema derivado de um problema MPEC para um ponto viável arbitrário x^* . Dada uma partição $(\beta_1, \beta_2) \in \mathcal{P}(I_{00}) =$

$\{(\beta_1, \beta_2) : \beta_1 \cup \beta_2 = I_{00}, \beta_1 \cap \beta_2 = \emptyset\}$ o problema

$$\begin{aligned} \min \quad & f(x) \\ \text{s.a.} \quad & g(x) \leq 0, \quad h(x) = 0, \\ & G_{I_{0+} \cup \beta_1}(x) = 0, \quad H_{I_{0+} \cup \beta_1}(x) \geq 0, \\ & H_{I_{+0} \cup \beta_2}(x) = 0, \quad G_{I_{+0} \cup \beta_2}(x) \geq 0 \end{aligned} \quad (1.14)$$

é chamado de *problema não linear* (β_1, β_2) -restrito [58] e é denotado por $NLP_*(\beta_1, \beta_2)$ (depende de x^*). Seu cone tangente associado e cone viável linearizado em x^* são denotados por

$$\mathcal{T}_{NLP_*(\beta_1, \beta_2)}(x^*) \quad \text{e} \quad \mathcal{T}_{NLP_*(\beta_1, \beta_2)}^{lin}(x^*),$$

respectivamente, tendo o último a seguinte forma:

$$\begin{aligned} \mathcal{T}_{NLP_*(\beta_1, \beta_2)}^{lin}(x^*) = \{d \in \mathbb{R}^n \mid & \nabla g_i(x^*)^T d \leq 0, \quad \forall i \in \mathcal{I}_g \\ & \nabla h_i(x^*)^T d = 0, \quad \forall i = 1, \dots, p \\ & \nabla G_i(x^*)^T d = 0, \quad \forall i \in I_{0+} \cup \beta_1 \\ & \nabla H_i(x^*)^T d = 0, \quad \forall i \in I_{+0} \cup \beta_2 \\ & \nabla G_i(x^*)^T d \geq 0, \quad \forall i \in \beta_2 \\ & \nabla H_i(x^*)^T d \geq 0, \quad \forall i \in \beta_1\}. \end{aligned} \quad (1.15)$$

Lema 1.18. [21, Lema 3.1] *Seja x^* um ponto viável do Problema MPEC (1). Então*

$$\mathcal{T}(x^*) = \bigcup_{(\beta_1, \beta_2) \in \mathcal{P}(I_{00})} \mathcal{T}_{NLP_*(\beta_1, \beta_2)}(x^*).$$

Demonstração. Seja \mathcal{X} o conjunto viável do Problema MPEC (1) e $\mathcal{X}_{(\beta_1, \beta_2)}$ o conjunto viável do problema não linear (β_1, β_2) -restrito $NLP_*(\beta_1, \beta_2)$ (1.14). Mostremos primeiro que existe uma vizinhança $\mathcal{N} \in \mathbb{R}^n$ de x^* tal que

$$\mathcal{X} \cap \mathcal{N} = \bigcup_{(\beta_1, \beta_2) \in \mathcal{P}(I_{00})} (\mathcal{X}_{(\beta_1, \beta_2)} \cap \mathcal{N}). \quad (1.16)$$

De fato, comparando ambos problemas:

$\begin{aligned} \min \quad & f(x) \\ \text{s.a.} \quad & g(x) \leq 0, \quad h(x) = 0 \\ & G(x) \geq 0, \quad H(x) \geq 0, \\ & G(x)^T H(x) = 0 \end{aligned}$	$\begin{aligned} \min \quad & f(x) \\ \text{s.a.} \quad & g(x) \leq 0, \quad h(x) = 0 \\ & G_{I_{0+} \cup \beta_1}(x) = 0, \quad H_{I_{0+} \cup \beta_1}(x) \geq 0, \\ & H_{I_{+0} \cup \beta_2}(x) = 0, \quad G_{I_{+0} \cup \beta_2}(x) \geq 0 \end{aligned}$
--	---

Notamos primeiro que as restrições $g(x) \leq 0$ e $h(x) = 0$ aparecem em \mathcal{X} e $\mathcal{X}_{(\beta_1, \beta_2)}$ logo podem ser ignoradas. Além disso como $\mathcal{X}_{(\beta_1, \beta_2)} \subseteq \mathcal{X}$ para todo $(\beta_1, \beta_2) \in \mathcal{P}(I_{00})$, a inclusão “ \supseteq ” em (1.16) é imediata, ficando só a parte da “ \subseteq ”. Consideremos i arbitrário

tal que $i \in I_{0+}$ ($G_i(x^*) = 0$ e $H_i(x^*) > 0$), logo como H é contínua existe $\varepsilon_i > 0$ tal que $H_i(x) > 0$ para todo $x \in B_{\varepsilon_i}(x^*)$ e pela complementaridade $G_i(x) = 0$ para todo $x \in B_{\varepsilon_i}(x^*) \cap \mathcal{X}$. Analogamente para todo $i \in I_{+0}$ também podemos encontrar $\varepsilon_i > 0$ tal que $G_i(x) > 0$ para todo $x \in B_{\varepsilon_i}(x^*)$ e $H_i(x) = 0$ para todo $x \in B_{\varepsilon_i}(x^*) \cap \mathcal{X}$. Logo considerando $\varepsilon = \min_{i \in I_{0+} \cup I_{+0}} \{\varepsilon_i\}$ temos que se $x \in B_\varepsilon(x^*) \cap \mathcal{X}$, então

$$\begin{aligned} G_{I_{0+}}(x) &= 0 \text{ e } H_{I_{0+}}(x) > 0 \quad e \\ H_{I_{+0}}(x) &= 0 \text{ e } G_{I_{+0}}(x) > 0. \end{aligned}$$

Para $x \in B_\varepsilon(x^*) \cap \mathcal{X}$, defina $\beta_1 = \{i \in I_{00} \mid G_i(x) = 0\}$ e $\beta_2 = I_{00} \setminus \beta_1$, então pela complementaridade $H_{\beta_2}(x) = 0$. Com este resultado e considerando $\mathcal{N} = B_\varepsilon(x^*)$, temos que se $x \in \mathcal{X} \cap \mathcal{N}$, então a partição $(\beta_1, \beta_2) \in \mathcal{P}(I_{00})$ satisfaz

$$\begin{aligned} G_{I_{0+} \cup \beta_1}(x) &= 0 \text{ e } H_{I_{0+} \cup \beta_1}(x) \geq 0 \quad e \\ H_{I_{+0} \cup \beta_2}(x) &= 0 \text{ e } G_{I_{+0} \cup \beta_2}(x) \geq 0, \end{aligned}$$

isto é, $x \in \mathcal{X}_{(\beta_1, \beta_2)} \cap \mathcal{N}$, concluindo (1.16).

Finalmente, temos

$$\begin{aligned} \mathcal{T}(x^*) &= \mathcal{T}(x^*, \mathcal{X}) \\ &= \mathcal{T}(x^*, \mathcal{X} \cap \mathcal{N}) \\ &= \mathcal{T}(x^*, \bigcup_{(\beta_1, \beta_2) \in \mathcal{P}(I_{00})} (\mathcal{X}_{(\beta_1, \beta_2)} \cap \mathcal{N})) \\ &= \bigcup_{(\beta_1, \beta_2) \in \mathcal{P}(I_{00})} \mathcal{T}(x^*, (\mathcal{X}_{(\beta_1, \beta_2)} \cap \mathcal{N})) \\ &= \bigcup_{(\beta_1, \beta_2) \in \mathcal{P}(I_{00})} \mathcal{T}(x^*, \mathcal{X}_{(\beta_1, \beta_2)}) \\ &= \bigcup_{(\beta_1, \beta_2) \in \mathcal{P}(I_{00})} \mathcal{T}_{NLP_*(\beta_1, \beta_2)}(x^*). \end{aligned}$$

□

Lema 1.19. [58, Proposição 3] *Seja x^* um ponto viável do Problema MPEC (1). Se ACQ é válido para todo problema $NLP_*(\beta_1, \beta_2)$, isto é, se*

$$\mathcal{T}_{NLP_*(\beta_1, \beta_2)}(x^*) = \mathcal{T}_{NLP_*(\beta_1, \beta_2)}^{lin}(x^*) \quad (1.17)$$

para todo $(\beta_1, \beta_2) \in \mathcal{P}(I_{00})$, então $\mathcal{T}(x^*) = \mathcal{K}(x^*)$ sendo

$$\begin{aligned} \mathcal{K}(x^*) = \{d \in \mathbb{R}^n \mid & \nabla g_i(x^*)^T d \leq 0, \quad \forall i \in \mathcal{I}_g \\ & \nabla h_i(x^*)^T d = 0, \quad \forall i = 1, \dots, p \\ & \nabla G_i(x^*)^T d = 0, \quad \forall i \in I_{0+} \\ & \nabla H_i(x^*)^T d = 0, \quad \forall i \in I_{+0} \\ & \nabla G_i(x^*)^T d \geq 0, \quad \forall i \in I_{00} \\ & \nabla H_i(x^*)^T d \geq 0, \quad \forall i \in I_{00} \\ & (\nabla G_i(x^*)^T d) (\nabla H_i(x^*)^T d) = 0, \quad \forall i \in I_{00}\}. \end{aligned} \quad (1.18)$$

Demonstração. Seja $d \in \mathcal{T}(x^*)$, então pelo Lema 1.18 temos que existe uma partição $(\beta_1, \beta_2) \in \mathcal{P}(I_{00})$ tal que $d \in \mathcal{T}_{NLP_*(\beta_1, \beta_2)}(x^*)$. Logo pela hipótese (1.17) temos que d satisfaz (1.15) de onde

$$(\nabla G_i(x^*)^T d) (\nabla H_i(x^*)^T d) = 0 \quad \forall i \in I_{00} \quad (1.19)$$

$$\nabla G_i(x^*)^T d \geq 0 \text{ e } \nabla H_i(x^*)^T d \geq 0 \quad \forall i \in I_{00}, \quad (1.20)$$

provando assim que $d \in \mathcal{K}(x^*)$. Para provar a inclusão contrária, considere $d \in \mathcal{K}(x^*)$ e defina $\beta_1 = \{i \in I_{00} \mid \nabla G_i(x^*)^T d = 0\}$ e $\beta_2 = I_{00} \setminus \beta_1$ então pela complementaridade em $\mathcal{K}(x^*)$, $\nabla H_{\beta_2}(x^*)^T d = 0$, logo $d \in \mathcal{T}_{NLP_*(\beta_1, \beta_2)}^{lin}(x^*) = \mathcal{T}_{NLP_*(\beta_1, \beta_2)}(x^*)$ e pelo Lema 1.18 $d \in \mathcal{T}(x^*)$. \square

Proposição 1.20. *Seja x^* um ponto viável do Problema MPEC (1). Suponha que ACQ é válido para todo problema $NLP_*(\beta_1, \beta_2)$, isto é,*

$$\mathcal{T}_{NLP_*(\beta_1, \beta_2)}(x^*) = \mathcal{T}_{NLP_*(\beta_1, \beta_2)}^{lin}(x^*) \quad (1.21)$$

para todo $(\beta_1, \beta_2) \in \mathcal{P}(I_{00})$. Se x^* satisfaz ACQ para o Problema MPEC (1), então existe uma partição $(\hat{\beta}_1, \hat{\beta}_2) \in \mathcal{P}(I_{00})$ tal que $\mathcal{T}(x^*) = \mathcal{T}_{NLP_*(\hat{\beta}_1, \hat{\beta}_2)}(x^*)$.

Demonstração. Primeiramente note que pelo Lema 1.19 temos que $\mathcal{T}(x^*)$ é dado por (1.18). Considere a partição $(\hat{\beta}_1, \hat{\beta}_2) \in \mathcal{P}(I_{00})$ definida por

$$\hat{\beta}_1 = \{i \in I_{00} \mid \nabla G_i(x^*)^T d = 0 \text{ para todo } d \in \mathcal{T}(x^*)\} \quad \text{e} \quad \hat{\beta}_2 = I_{00} \setminus \hat{\beta}_1.$$

Afirmamos que

$$\mathcal{T}(x^*) \subset \mathcal{T}_{NLP_*(\hat{\beta}_1, \hat{\beta}_2)}(x^*). \quad (1.22)$$

De fato, dado $d \in \mathcal{T}(x^*)$ arbitrário, comparando ambos os cones

$\mathcal{T}(x^*)$	$\mathcal{T}_{NLP_*(\hat{\beta}_1, \hat{\beta}_2)}^{lin}(x^*)$
$\{d \in \mathbb{R}^n \mid \nabla g_i(x^*)^T d \leq 0, \forall i \in \mathcal{I}_g$ $\nabla h_i(x^*)^T d = 0, \forall i = 1, \dots, p$ $\nabla G_i(x^*)^T d = 0, \forall i \in I_{0+}$ $\nabla H_i(x^*)^T d = 0, \forall i \in I_{+0}$ $\nabla G_i(x^*)^T d \geq 0, \forall i \in I_{00}$ $\nabla H_i(x^*)^T d \geq 0, \forall i \in I_{00}$ $(\nabla G_i(x^*)^T d) (\nabla H_i(x^*)^T d) = 0, \forall i \in I_{00}\}$	$\{d \in \mathbb{R}^n \mid \nabla g_i(x^*)^T d \leq 0, \forall i \in \mathcal{I}_g$ $\nabla h_i(x^*)^T d = 0, \forall i = 1, \dots, p$ $\nabla G_i(x^*)^T d = 0, \forall i \in I_{0+} \cup \hat{\beta}_1$ $\nabla H_i(x^*)^T d = 0, \forall i \in I_{+0} \cup \hat{\beta}_2$ $\nabla G_i(x^*)^T d \geq 0, \forall i \in \hat{\beta}_2$ $\nabla H_i(x^*)^T d \geq 0, \forall i \in \hat{\beta}_1\}$

vemos que basta mostrar que $\nabla H_i(x^*)^T d = 0$ para todo $i \in \hat{\beta}_2$. Para tais índices temos que existe $\bar{d} \in \mathcal{T}(x^*)$ tal que $\nabla G_i(x^*)^T \bar{d} > 0$. Portanto, pela última relação em (1.18), temos que $\nabla H_i(x^*)^T \bar{d} = 0$. Por outro lado, como x^* satisfaz ACQ, temos pelo Lema 1.7 (a) que $\mathcal{T}(x^*)$ é convexo, logo $\tilde{d} = (1-t)d + t\bar{d} \in \mathcal{T}(x^*)$ para todo $t \in [0, 1]$. Usando novamente a última relação em (1.18), obtemos

$$\begin{aligned}
0 &= \left(\nabla G_i(x^*)^T \tilde{d} \right) \left(\nabla H_i(x^*)^T \tilde{d} \right) \\
&= (1-t)t \left(\left(\nabla G_i(x^*)^T d \right) \left(\nabla H_i(x^*)^T \bar{d} \right) + \left(\nabla G_i(x^*)^T \bar{d} \right) \left(\nabla H_i(x^*)^T d \right) \right) \\
&= (1-t)t \left(\nabla G_i(x^*)^T \bar{d} \right) \left(\nabla H_i(x^*)^T d \right),
\end{aligned}$$

donde segue que $\nabla H_i(x^*)^T d = 0$, provando (1.22). A inclusão contrária segue do Lema 1.18. Desta forma obtemos $\mathcal{T}(x^*) = \mathcal{T}_{NLP_*(\hat{\beta}_1, \hat{\beta}_2)}(x^*)$. \square

O resultado estabelecido na proposição anterior também é apresentado em [23, Proposição 2.1]. No entanto, em [23] também é feita a afirmação recíproca, com uma suposta demonstração. Porém, como veremos no exemplo seguinte esta afirmação não é verdadeira.

Exemplo 1.21. *Considere o problema*

$$\begin{aligned}
\min \quad & f(x) = x_1^2 + x_2^2 + x_3^2 \\
s.a. \quad & h(x) = x_3 = 0 \\
& G_1(x) = x_2 + x_3^2 \geq 0 \\
& H_1(x) = x_2 - x_3^2 \geq 0 \\
& G_2(x) = x_2 + x_3^3 \geq 0 \\
& H_2(x) = x_2 - x_3^3 \geq 0 \\
& \Theta(x) = G(x)^T H(x) = (x_2 + x_3^2)(x_2 - x_3^2) + (x_2 + x_3^3)(x_2 - x_3^3) = 0.
\end{aligned}$$

O conjunto viável é $\left\{ \begin{pmatrix} t & 0 & 0 \end{pmatrix}^T \mid t \in \mathbb{R} \right\}$ e $x^* = \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}^T$ o único minimizador. Claramente temos $\mathcal{T}(x^*) = \left\{ \begin{pmatrix} t & 0 & 0 \end{pmatrix}^T \mid t \in \mathbb{R} \right\}$. Além disso, como $\nabla h(x^*) =$

$\begin{pmatrix} 0 & 0 & 1 \end{pmatrix}^T$, $\nabla G_1(x^*) = \nabla H_1(x^*) = \nabla G_2(x^*) = \nabla H_2(x^*) = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}^T$ e $\nabla \Theta(x^*) = \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}^T$, então $\mathcal{T}^{lin}(x^*) = \left\{ \begin{pmatrix} t & s & 0 \end{pmatrix}^T \mid t \in \mathbb{R}, s \geq 0 \right\}$. Logo x^* não satisfaz ACQ. Por outro lado $I_{00} = \{1, 2\}$, $I_{0+} = \emptyset$, $I_{+0} = \emptyset$ e as únicas partições (β_1, β_2) de I_{00} são

$$\begin{aligned} \beta_1 &= \{1, 2\} \text{ e } \beta_2 = \emptyset, \\ \beta_2 &= \{1, 2\} \text{ e } \beta_1 = \emptyset, \\ \beta_1 &= \{1\} \text{ e } \beta_2 = \{2\} \text{ e} \\ \beta_2 &= \{1\} \text{ e } \beta_1 = \{2\}. \end{aligned}$$

Neste caso os problemas $NLP_*(\beta_1, \beta_2)$ têm a forma

$\begin{aligned} \min \quad & f(x) \\ \text{s.a.} \quad & h(x) = x_3 = 0 \\ & G_1(x) = x_2 + x_3^2 = 0 \\ & H_1(x) = x_2 - x_3^2 \geq 0 \\ & G_2(x) = x_2 + x_3^3 = 0 \\ & H_2(x) = x_2 - x_3^3 \geq 0, \end{aligned}$	$\begin{aligned} \min \quad & f(x) \\ \text{s.a.} \quad & h(x) = x_3 = 0 \\ & G_1(x) = x_2 + x_3^2 \geq 0 \\ & H_1(x) = x_2 - x_3^2 = 0 \\ & G_2(x) = x_2 + x_3^3 \geq 0 \\ & H_2(x) = x_2 - x_3^3 = 0, \end{aligned}$
$\begin{aligned} \min \quad & f(x) \\ \text{s.a.} \quad & h(x) = x_3 = 0 \\ & G_1(x) = x_2 + x_3^2 = 0 \\ & H_1(x) = x_2 - x_3^2 \geq 0 \\ & G_2(x) = x_2 + x_3^3 \geq 0 \\ & H_2(x) = x_2 - x_3^3 = 0, \end{aligned}$	$\begin{aligned} \min \quad & f(x) \\ \text{s.a.} \quad & h(x) = x_3 = 0 \\ & G_1(x) = x_2 + x_3^2 \geq 0 \\ & H_1(x) = x_2 - x_3^2 = 0 \\ & G_2(x) = x_2 + x_3^3 = 0 \\ & H_2(x) = x_2 - x_3^3 \geq 0, \end{aligned}$

sendo o conjunto viável $\left\{ \begin{pmatrix} t & 0 & 0 \end{pmatrix}^T \mid t \in \mathbb{R} \right\}$ e $\mathcal{T}_{NLP_*(\beta_1, \beta_2)}(x^*) = \mathcal{T}_{NLP_*(\beta_1, \beta_2)}^{lin}(x^*) = \left\{ \begin{pmatrix} t & 0 & 0 \end{pmatrix}^T \mid t \in \mathbb{R} \right\}$ para todos os casos, isto é x^* satisfaz ACQ para os problemas $NLP_*(\beta_1, \beta_2)$.

Cabe ressaltar que não é usual que ACQ seja satisfeita em problemas MPEC, como se pode ver neste exemplo simples.

Exemplo 1.22. Considere o problema

$$\begin{aligned} \min \quad & x_1^2 + x_2^2 \\ \text{s.a.} \quad & x_1 \geq 0 \\ & x_2 \geq 0 \\ & x_1 x_2 = 0. \end{aligned}$$

Neste caso $x^* = \begin{pmatrix} 0 & 0 \end{pmatrix}^T$ é o único minimizador, $\mathcal{T}(x^*) = \{x \mid x_1 \geq 0, x_2 \geq 0 \text{ e } x_1 x_2 = 0\}$ e $\mathcal{T}^{lin}(x^*) = \{x \mid x_1 \geq 0, x_2 \geq 0\}$. Logo x^* não satisfaz ACQ. Porém

$\mathcal{T}(x^*)^* = \mathcal{T}^{lin}(x^*)^* = \{x \mid x_1 \leq 0, x_2 \leq 0\}$, isto é, x^* satisfaz GCQ. Na Seção 2.2 veremos que esta condição é mais usual de ser satisfeita. Antes disso, veremos no próximo capítulo outras condições de qualificação associadas a problemas MPEC.

Capítulo 2

Condições de qualificação e estacionariedade em MPEC

Para poder identificar um minimizador local do Problema MPEC (1), vários conceitos de estacionariedade foram introduzidos ao longo dos anos, no entanto, as restrições do problema devem satisfazer condições de qualificação (CQs) para que um minimizador local satisfaça algum conceito de estacionariedade. Neste capítulo, lidaremos com a definição de CQs e conceitos de estacionariedade associados a problemas MPEC, descrevendo a relação que existe entre esses conceitos, bem como sua relação com as CQs usuais para problemas não lineares em geral.

2.1 Condições de qualificação MPEC usuais

Como vimos no capítulo anterior, algumas das CQs mais comuns na literatura não são satisfeitas por ponto viável algum dos Problemas MPEC (1). Para superar isso, são introduzidas novas condições de qualificação específicas para problemas MPEC. Tendo em mente as quatro CQs vistas na Seção 1.1, nesta seção nos ocuparemos das CQs MPEC associadas a elas. Descreveremos a relação que existe entre estes novos conceitos e concluiremos com uma relação similar a (1.9).

2.1.1 MPEC-LICQ e MPEC-MFCQ

Seja x^* um ponto viável do Problema MPEC (1), consideremos o seguinte problema, chamado *MPEC reduzido*

$$\begin{aligned} \min \quad & f(x) \\ \text{s.a} \quad & g(x) \leq 0, \quad h(x) = 0 \\ & G_{I_{0+} \cup I_{00}}(x) = 0, \quad G_{I_{+0}}(x) \geq 0 \\ & H_{I_{+0} \cup I_{00}}(x) = 0, \quad H_{I_{0+}}(x) \geq 0. \end{aligned} \tag{2.1}$$

O problema anterior é chamado reduzido uma vez que sua região viável é um subconjunto da região viável de (1). Note também que se x^* é um minimizador local do Problema MPEC (1) e ao mesmo tempo é viável para o Problema (2.1), então x^* é um minimizador local do problema do correspondente reduzido (2.1).

Definição 2.1. *Seja x^* um ponto viável do Problema MPEC (1). A condição MPEC-LICQ (MPEC-MFCQ) é satisfeita em x^* se o correspondente Problema reduzido (2.1) satisfaz LICQ (MFCQ) em x^* .*

Note que pela definição MPEC-LICQ implica MPEC-MFCQ, como no caso das CQs usuais.

2.1.2 Cone Viável Linearizado MPEC

Lembremos que a CQ de Abadie, $\mathcal{T}(x^*) = \mathcal{T}^{lin}(x^*)$, obriga ao cone tangente $\mathcal{T}(x^*)$ ser convexo (Lema 1.7 (a)), fato que em geral não é comum em problemas MPEC como pode ser visto no Exemplo 1.22. Numa forma de abordar este inconveniente é encontrar outro cone, não necessariamente convexo, “ X ” que cumpra a relação

$$\mathcal{T}(x^*) \subseteq X \subseteq \mathcal{T}^{lin}(x^*),$$

e que permita estabelecer condições semelhantes às de Abadie e Guignard.

Definição 2.2. [21] *Seja x^* um ponto viável do Problema MPEC (1). Definimos o cone viável linearizado MPEC em x^* como*

$$\begin{aligned} \mathcal{T}_{MPEC}^{lin}(x^*) = \{d \in \mathbb{R}^n \mid & \nabla g_i(x^*)^T d \leq 0, \quad \forall i \in \mathcal{I}_g \\ & \nabla h_i(x^*)^T d = 0, \quad \forall i = 1, \dots, p \\ & \nabla G_i(x^*)^T d = 0, \quad \forall i \in I_{0+} \\ & \nabla H_i(x^*)^T d = 0, \quad \forall i \in I_{+0} \\ & \nabla G_i(x^*)^T d \geq 0, \quad \forall i \in I_{00} \\ & \nabla H_i(x^*)^T d \geq 0, \quad \forall i \in I_{00} \\ & (\nabla G_i(x^*)^T d) (\nabla H_i(x^*)^T d) = 0, \quad \forall i \in I_{00}\}. \end{aligned} \tag{2.2}$$

Comparando este cone com o cone $\mathcal{T}^{lin}(x^*)$, que no caso de problemas MPEC tem a forma (1.13), temos claramente que

$$\mathcal{T}_{MPEC}^{lin}(x^*) \subseteq \mathcal{T}^{lin}(x^*). \tag{2.3}$$

O cone viável linearizado MPEC (2.2) também satisfaz $\mathcal{T}(x^*) \subseteq \mathcal{T}_{MPEC}^{lin}(x^*)$. Para provar esta inclusão veremos primeiramente o seguinte resultado:

Lema 2.3. [21, Lema 3.1] Se x^* é um ponto viável do Problema MPEC (1) então

$$\mathcal{T}_{MPEC}^{lin}(x^*) = \bigcup_{(\beta_1, \beta_2) \in \mathcal{P}(I_{00})} \mathcal{T}_{NLP_*(\beta_1, \beta_2)}^{lin}(x^*).$$

Demonstração. Temos que

$\mathcal{T}_{MPEC}^{lin}(x^*)$	$\mathcal{T}_{NLP_*(\beta_1, \beta_2)}^{lin}(x^*)$
$\{d \in \mathbb{R}^n \mid \nabla g_i(x^*)^T d \leq 0, \forall i \in \mathcal{I}_g$	$\{d \in \mathbb{R}^n \mid \nabla g_i(x^*)^T d \leq 0, \forall i \in \mathcal{I}_g$
$\nabla h_i(x^*)^T d = 0, \forall i = 1, \dots, p$	$\nabla h_i(x^*)^T d = 0, \forall i = 1, \dots, p$
$\nabla G_i(x^*)^T d = 0, \forall i \in I_{0+}$	$\nabla G_i(x^*)^T d = 0, \forall i \in I_{0+} \cup \beta_1$
$\nabla H_i(x^*)^T d = 0, \forall i \in I_{+0}$	$\nabla H_i(x^*)^T d = 0, \forall i \in I_{+0} \cup \beta_2$
$\nabla G_i(x^*)^T d \geq 0, \forall i \in I_{00}$	$\nabla G_i(x^*)^T d \geq 0, \forall i \in \beta_2$
$\nabla H_i(x^*)^T d \geq 0, \forall i \in I_{00}$	$\nabla H_i(x^*)^T d \geq 0, \forall i \in \beta_1\}$
$(\nabla G_i(x^*)^T d) (\nabla H_i(x^*)^T d) = 0, \forall i \in I_{00}\}$	

Para $d \in \mathcal{T}_{MPEC}^{lin}(x^*)$ fixo, considere

$$\beta_1 = \{i \in I_{00} : \nabla G_i(x^*)^T d = 0\} \text{ e}$$

$$\beta_2 = \{i \in I_{00} : \nabla G_i(x^*)^T d > 0\}$$

então pela complementaridade em $\mathcal{T}_{MPEC}^{lin}(x^*)$,

$$\nabla H_i(x^*)^T d = 0, \forall i \in \beta_2$$

e em consequência $d \in \mathcal{T}_{NLP_*(\beta_1, \beta_2)}^{lin}(x^*)$.

Para a inclusão contrária, considere uma partição arbitrária $(\beta_1, \beta_2) \in \mathcal{P}(I_{00})$. Temos da definição de $\mathcal{T}_{NLP_*(\beta_1, \beta_2)}^{lin}(x^*)$,

$$(\nabla G_i(x^*)^T d) (\nabla H_i(x^*)^T d) = 0, \forall i \in I_{00}$$

de onde

$$\mathcal{T}_{NLP_*(\beta_1, \beta_2)}^{lin}(x^*) \subseteq \mathcal{T}_{MPEC}^{lin}(x^*).$$

□

Com este lema podemos agora encontrar a relação entre $\mathcal{T}(x^*)$ e $\mathcal{T}_{MPEC}^{lin}(x^*)$.

Lema 2.4. [21, Lema 3.2] Se x^* é um ponto viável do Problema MPEC (1) então

$$\mathcal{T}(x^*) \subseteq \mathcal{T}_{MPEC}^{lin}(x^*).$$

Demonstração. Como o cone tangente está sempre incluído no cone viável linearizado,

temos para o caso do Problema não linear (β_1, β_2) -restrito $NLP_*(\beta_1, \beta_2)$ (1.14) que

$$\mathcal{T}_{NLP_*(\beta_1, \beta_2)}(x^*) \subseteq \mathcal{T}_{NLP_*(\beta_1, \beta_2)}^{lin}(x^*)$$

e portanto

$$\bigcup_{(\beta_1, \beta_2) \in \mathcal{P}(I_{00})} \mathcal{T}_{NLP_*(\beta_1, \beta_2)}(x^*) \subseteq \bigcup_{(\beta_1, \beta_2) \in \mathcal{P}(I_{00})} \mathcal{T}_{NLP_*(\beta_1, \beta_2)}^{lin}(x^*).$$

Então pelos Lemas 1.18 e 2.3,

$$\mathcal{T}(x^*) \subseteq \mathcal{T}_{MPEC}^{lin}(x^*).$$

□

Em resumo, pela relação (2.3) e o Lema 2.4, temos

$$\mathcal{T}(x^*) \subseteq \mathcal{T}_{MPEC}^{lin}(x^*) \subseteq \mathcal{T}^{lin}(x^*).$$

2.1.3 MPEC-ACQ e MPEC-GCQ

No caso da programação não-linear usual, as CQs de Abadie e Guignard estão baseadas na relação que existe entre o cone tangente $\mathcal{T}(x^*)$ e o cone viável linearizado $\mathcal{T}^{lin}(x^*)$. Assim as CQs de Abadie e Guignard estão dadas por $\mathcal{T}(x^*) = \mathcal{T}^{lin}(x^*)$ e $\mathcal{T}(x^*)^* = \mathcal{T}^{lin}(x^*)^*$ respectivamente. No caso de problemas MPEC, elas são definidas de forma semelhante, considerando o cone $\mathcal{T}_{MPEC}^{lin}(x^*)$ (2.2) em vez do cone viável linearizado $\mathcal{T}^{lin}(x^*)$.

Definição 2.5. [22, Definição 2.1] *Seja x^* um ponto viável do Problema MPEC (1). Dizemos que x^* satisfaz a condição de qualificação MPEC-Abadie (MPEC-ACQ) se*

$$\mathcal{T}(x^*) = \mathcal{T}_{MPEC}^{lin}(x^*),$$

e satisfaz a condição de qualificação MPEC-Guignard (MPEC-GCQ) se

$$\mathcal{T}(x^*)^* = \mathcal{T}_{MPEC}^{lin}(x^*)^*.$$

É claro que MPEC-ACQ implica MPEC-GCQ e, como no caso das CQ usuais, a recíproca não é verdadeira, como pode ser visto no exemplo a seguir:

Exemplo 2.6. [59] *Considere o seguinte problema*

$$\begin{aligned} \min \quad & x_1^2 + x_2^2 \\ \text{s.a} \quad & G(x) = x_1^2 \geq 0 \\ & H(x) = x_2^2 \geq 0 \\ & G(x)^T H(x) = x_1^2 x_2^2 = 0, \end{aligned}$$

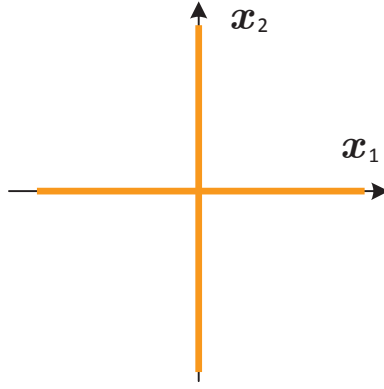


Figura 2.1: Conjunto viável do Exemplo 2.6.

Temos que seu conjunto viável está formado pelos dos eixos coordenados (Figura 2.1) e evidentemente a origem $x^* = (0 \ 0)^T$ é seu único minimizador com $\nabla G(x^*) = \nabla H(x^*) = (0 \ 0)^T$. Assim $I_{0+} = I_{+0} = \emptyset$ e $I_{00} = \{1\}$ e neste caso o cone viável linearizado MPEC (2.2), é dado por

$$\begin{aligned} \mathcal{T}_{MPEC}^{lin}(x^*) = \{d \in \mathbb{R}^2 \mid & \nabla G(x^*)^T d \geq 0, \\ & \nabla H(x^*)^T d \geq 0, \\ & (\nabla G(x^*)^T d) (\nabla H(x^*)^T d) = 0\}, \end{aligned}$$

ou seja, $\mathcal{T}_{MPEC}^{lin}(x^*) = \mathbb{R}^2$. Por outro lado, podemos ver que o cone tangente é

$$\mathcal{T}(x^*) = \{d \in \mathbb{R}^2 \mid d_1 d_2 = 0\},$$

resultando que

$$\mathcal{T}(x^*) \neq \mathcal{T}_{MPEC}^{lin}(x^*),$$

porém, considerando os duais, temos

$$\mathcal{T}(x^*)^* = \mathcal{T}_{MPEC}^{lin}(x^*)^* = \begin{pmatrix} 0 & 0 \end{pmatrix}^T,$$

e assim MPEC-GCQ é mais fraca que MPEC-ACQ.

2.1.4 MPEC-ACQ e MPEC-MFCQ

Nas seções anteriores, vimos as relações entre MPEC-LICQ e MPEC-MFCQ e entre MPEC-ACQ e MPEC-GCQ. Para fechar esta seção, podemos provar o seguinte resultado, análogo ao padrão em programação não-linear, que mostra que MPEC-MFCQ implica MPEC-ACQ. Para provar isso, veremos alguns resultados auxiliares.

Lema 2.7. *Seja x^* um ponto viável do Problema MPEC (1). Se x^* satisfaz ACQ para o Problema $NPL_*(\beta_1, \beta_2)$ (1.14) para toda partição $(\beta_1, \beta_2) \in \mathcal{P}(I_{00})$, então x^* satisfaz*

MPEC-ACQ para o Problema MPEC (1).

Demonstração. Como para toda partição $(\beta_1, \beta_2) \in \mathcal{P}(I_{00})$ x^* satisfaz ACQ para o Problema $NLP_*(\beta_1, \beta_2)$ (1.14), então o cone tangente e o cone viável linearizado são iguais em x^* (Definição 1.10), ou seja

$$\mathcal{T}_{NLP_*(\beta_1, \beta_2)}(x^*) = \mathcal{T}_{NLP_*(\beta_1, \beta_2)}^{lin}(x^*).$$

Portanto

$$\bigcup_{(\beta_1, \beta_2) \in \mathcal{P}(I_{00})} \mathcal{T}_{NLP_*(\beta_1, \beta_2)}(x^*) = \bigcup_{(\beta_1, \beta_2) \in \mathcal{P}(I_{00})} \mathcal{T}_{NLP_*(\beta_1, \beta_2)}^{lin}(x^*),$$

que em vista dos Lemas 1.18 e 2.3, nos permite concluir que

$$\mathcal{T}(x^*) = \mathcal{T}_{MPEC}^{lin}(x^*),$$

o que prova o lema. □

Lema 2.8. *Se $\{a_i \in \mathbb{R}^n \mid i = 1, \dots, m\}$ é um conjunto linearmente independente e existe um vetor não nulo $d \in \mathbb{R}^n$, tal que*

$$a_i^T d = 0, \forall i = 1, \dots, m. \quad (2.4)$$

Então, para todo $r \leq m$, existe um vetor $\hat{d} \in \mathbb{R}^n$ tal que

$$\begin{aligned} a_i^T (d + \delta \hat{d}) &= 0, \forall i = 1, \dots, r \\ a_i^T (d + \delta \hat{d}) &> 0, \forall i = r + 1, \dots, m. \end{aligned}$$

para todo $\delta > 0$.

Demonstração. Como $\{a_i \in \mathbb{R}^n \mid i = 1, \dots, m\}$ é um conjunto linearmente independente, existe $\hat{d} \in \mathbb{R}^n$ tal que

$$\begin{aligned} a_i^T \hat{d} &= 0, \forall i = 1, \dots, r \\ a_i^T \hat{d} &= 1, \forall i = r + 1, \dots, m. \end{aligned}$$

Se $\delta > 0$ temos que

$$\begin{aligned} a_i^T (d + \delta \hat{d}) &= a_i^T d + \delta a_i^T \hat{d} = 0, \forall i = 1, \dots, r \\ a_i^T (d + \delta \hat{d}) &= a_i^T d + \delta a_i^T \hat{d} = \delta > 0, \forall i = r + 1, \dots, m. \end{aligned}$$

□

Lema 2.9. [21, Lema 3.3] *Se um ponto viável x^* do Problema MPEC (1) satisfaz MPEC-MFCQ, então x^* satisfaz MFCQ para o Problema $NLP_*(\beta_1, \beta_2)$ (1.14) para toda partição $(\beta_1, \beta_2) \in \mathcal{P}(I_{00})$.*

Demonstração. Temos pela definição de MPEC-MFCQ que os gradientes

$$\begin{aligned} \nabla h_i(x^*), i &= 1, \dots, p \\ \nabla G_i(x^*), i &\in I_{0+} \cup I_{00} \\ \nabla H_i(x^*), i &\in I_{+0} \cup I_{00} \end{aligned} \quad (2.5)$$

são linearmente independentes e existe um vetor $d \in \mathbb{R}^n$ tal que

$$\begin{aligned} \nabla g_i(x^*)^T d &< 0, \quad \forall i \in \mathcal{I}_g \\ \nabla h_i(x^*)^T d &= 0, \quad \forall i = 1, \dots, p \\ \nabla G_i(x^*)^T d &= 0, \quad \forall i \in I_{0+} \cup I_{00} \\ \nabla H_i(x^*)^T d &= 0, \quad \forall i \in I_{+0} \cup I_{00}. \end{aligned} \quad (2.6)$$

Lembrando que o Problema $NLP_*(\beta_1, \beta_2)$ (1.14) está definido por

$$\begin{aligned} \min \quad & f(x) \\ \text{s.a.} \quad & g(x) \leq 0, \quad h(x) = 0 \\ & G_{I_{0+} \cup \beta_1}(x) = 0, \quad H_{I_{0+} \cup \beta_1}(x) \geq 0, \\ & H_{I_{+0} \cup \beta_2}(x) = 0, \quad G_{I_{+0} \cup \beta_2}(x) \geq 0, \end{aligned}$$

temos por (2.5) que os gradientes das restrições de igualdade são linearmente independentes. Por outro lado, como os gradientes em (2.5) são linearmente independentes, temos pelo Lema 2.8 e por (2.6) que existe um vetor \hat{d} tal que para todo $\delta > 0$ e $d_1 = d + \delta \hat{d}$ vale

$$\begin{aligned} \nabla h_i(x^*)^T d_1 &= 0, \quad \forall i = 1, \dots, p \\ \nabla G_i(x^*)^T d_1 &= 0, \quad \forall i \in I_{0+} \cup \beta_1 \\ \nabla H_i(x^*)^T d_1 &= 0, \quad \forall i \in I_{+0} \cup \beta_2 \\ \nabla G_i(x^*)^T d_1 &> 0, \quad \forall i \in \beta_2 \\ \nabla H_i(x^*)^T d_1 &> 0, \quad \forall i \in \beta_1 \end{aligned}$$

e considerando δ suficientemente pequeno temos

$$\nabla g_i(x^*)^T d_1 < 0, \quad \forall i \in \mathcal{I}_g$$

assim x^* satisfaz MFCQ para o Problema $NLP_*(\beta_1, \beta_2)$ (1.14). □

Teorema 2.10. [21, Teorema 3.1] *Se um ponto viável x^* do Problema MPEC (1) satisfaz*

MPEC-MFCQ, então x^* satisfaz MPEC-ACQ.

Demonstração. Temos pelo Lema 2.9 que x^* satisfaz MFCQ para o problema $NLP_*(\beta_1, \beta_2)$ para toda partição $(\beta_1, \beta_2) \in \mathcal{P}(I_{00})$. Logo x^* satisfaz ACQ para o problema $NLP_*(\beta_1, \beta_2)$ para toda partição $(\beta_1, \beta_2) \in \mathcal{P}(I_{00})$. Finalmente, pelo Lema 2.7 x^* satisfaz MPEC-ACQ par o Problema MPEC (1). \square

Como no caso das CQs usuais, a recíproca do teorema anterior não é verdadeira, como pode ser visto no exemplo a seguir:

Exemplo 2.11. [58] Considere o seguinte problema:

$$\begin{aligned} \min \quad & x_1 - x_2 \\ \text{s.a} \quad & g(x) = x_2 \leq 0 \\ & G(x) = x_1 \geq 0 \\ & H(x) = x_1 + x_2 \geq 0 \\ & G(x)^T H(x) = x_1(x_1 + x_2) = 0. \end{aligned}$$

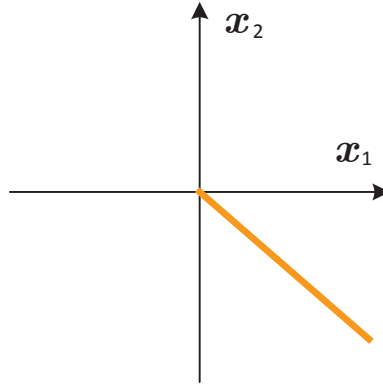


Figura 2.2: Conjunto viável do Exemplo 2.11.

O conjunto viável pode ser visto na Figura 2.2. Neste caso a origem $x^* = \begin{pmatrix} 0 & 0 \end{pmatrix}^T$ é seu único minimizador com $\nabla g(x^*) = \begin{pmatrix} 0 & 1 \end{pmatrix}^T$, $\nabla G(x^*) = \begin{pmatrix} 1 & 0 \end{pmatrix}^T$, $\nabla H(x^*) = \begin{pmatrix} 1 & 1 \end{pmatrix}^T$. Assim $I_{0+} = I_{+0} = \emptyset$, $I_{00} = \{1\}$. Como neste caso o cone viável linearizado MPEC (2.2) é dado por

$$\begin{aligned} \mathcal{T}_{MPEC}^{lin}(x^*) = \{d \in \mathbb{R}^2 \mid & \nabla g(x^*)^T d \leq 0 \\ & \nabla G(x^*)^T d \geq 0, \\ & \nabla H(x^*)^T d \geq 0, \\ & (\nabla G(x^*)^T d) (\nabla H(x^*)^T d) = 0\}, \end{aligned}$$

temos que se $d = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}$, então

$$\begin{aligned} \nabla g(x^*)^T d &= d_2 \leq 0 \\ \nabla G(x^*)^T d &= d_1 \geq 0 \\ \nabla H(x^*)^T d &= d_1 + d_2 \geq 0 \\ (\nabla G_i(x^*)^T d) (\nabla H_i(x^*)^T d) &= d_1 (d_1 + d_2) = 0. \end{aligned}$$

Agora se $d_1 + d_2 > 0$ então $d_1 = 0$ e $d_2 > 0$ o que contradiz a primeira linha; logo $d_1 + d_2 = 0$ e

$$\mathcal{T}_{MPEC}^{lin}(x^*) = \left\{ \begin{pmatrix} d_1 & d_2 \end{pmatrix}^T \in \mathbb{R}^2 : d_2 \leq 0, d_1 + d_2 = 0 \right\} = \mathcal{T}(x^*),$$

o que implica que vale MPEC-ACQ.

Por outro lado MPEC-MFCQ exige que exista $d = (d_1, d_2)^T \in \mathbb{R}^2$ tal que

$$\begin{aligned} d_2 &= \nabla g_i(x^*)^T d < 0, \quad \forall i \in \mathcal{I}_g \\ d_1 &= \nabla G_i(x^*)^T d = 0, \quad \forall i \in I_{0+} \cup I_{00} \\ d_1 + d_2 &= \nabla H_i(x^*)^T d = 0, \quad \forall i \in I_{+0} \cup I_{00}, \end{aligned}$$

assim x^* não satisfaz MPEC-MFCQ. Concluimos que MPEC-ACQ é mais fraca que MPEC-MFCQ.

Podemos resumir as relações entre as CQs MPEC estabelecidas nas seções anteriores através do diagrama a seguir:

$$\boxed{\text{MPEC-LICQ} \implies \text{MPEC-MFCQ} \implies \text{MPEC-ACQ} \implies \text{MPEC-GCQ}}$$

2.2 Relação entre as CQs MPEC e as CQs usuais

Evidentemente existem relações entre CQs MPEC e as CQs usuais para problemas não lineares em geral. A seguir veremos a relação de MPEC-LICQ e MPEC-ACQ com a mais fraca GCQ.

2.2.1 MPEC-LICQ e GCQ

Temos pela Definição 2.1 que se x^* satisfaz MPEC-LICQ para o Problema MPEC (1), então os gradientes

$$\begin{aligned} \nabla g_i(x^*), i &\in \mathcal{I}_g \\ \nabla h_i(x^*), i &= 1, \dots, p \\ \nabla G_i(x^*), i &\in I_{0+} \cup I_{00} \\ \nabla H_i(x^*), i &\in I_{+0} \cup I_{00} \end{aligned} \tag{2.7}$$

são linearmente independentes. Logo MPEC-LICQ é mais fraca que LICQ, pois neste caso, o conjunto de gradientes de (2.7) unido com o gradiente $\nabla (G(x^*)^T H(x^*))$ é linearmente independente; porém mesmo assim MPEC-LICQ implica GCQ.

Teorema 2.12. [23, Teorema 4.6] *Se um ponto viável x^* do Problema MPEC (1) satisfaz MPEC-LICQ, então satisfaz GCQ.*

Demonstração. Temos pelos Lemas 1.7 (b) e 1.8 (e) que $\mathcal{T}^{lin}(x^*)^* \subseteq \mathcal{T}(x^*)^*$, logo somente basta mostrar a inclusão contrária. Pelo Lema 1.18, temos que

$$\mathcal{T}(x^*) = \bigcup_{(\beta_1, \beta_2) \in \mathcal{P}(I_{00})} \mathcal{T}_{NLP_*(\beta_1, \beta_2)}(x^*)$$

o que implica pelo Lema 1.8 (d) que

$$\mathcal{T}(x^*)^* = \bigcap_{(\beta_1, \beta_2) \in \mathcal{P}(I_{00})} \mathcal{T}_{NLP_*(\beta_1, \beta_2)}(x^*)^*. \tag{2.8}$$

Por outro lado satisfazer MPEC-LICQ significa que os gradientes de (2.7) são linearmente independentes. Note que estas condições são as condições LICQ em x^* para o Problema $NLP_*(\beta_1, \beta_2)$ (1.14)

$$\begin{aligned} \min \quad & f(x) \\ \text{s.a} \quad & g(x) \leq 0, \quad h(x) = 0 \\ & G_{I_{0+} \cup \beta_1}(x) = 0, \quad H_{I_{0+} \cup \beta_1}(x) \geq 0, \\ & H_{I_{+0} \cup \beta_2}(x) = 0, \quad G_{I_{+0} \cup \beta_2}(x) \geq 0, \end{aligned}$$

para todo $(\beta_1, \beta_2) \in \mathcal{P}(I_{00})$. Logo x^* satisfaz GCQ para cada Problema $NLP_*(\beta_1, \beta_2)$ (1.14), o que implica que

$$\mathcal{T}_{NLP_*(\beta_1, \beta_2)}(x^*)^* = \mathcal{T}_{NLP_*(\beta_1, \beta_2)}^{lin}(x^*)^*,$$

para todo $(\beta_1, \beta_2) \in \mathcal{P}(I_{00})$. Além disso, por (1.15) o cone viável linearizado do problema

$NLP_*(\beta_1, \beta_2)$ tem a forma

$$\begin{aligned} \mathcal{T}_{NLP_*(\beta_1, \beta_2)}^{lin}(x^*) = \{ & d \in \mathbb{R}^n \mid \nabla g_i(x^*)^T d \leq 0, \quad \forall i \in \mathcal{I}_g \\ & \nabla h_i(x^*)^T d = 0, \quad \forall i = 1, \dots, p \\ & \nabla G_i(x^*)^T d = 0, \quad \forall i \in I_{0+} \cup \beta_1 \\ & \nabla H_i(x^*)^T d = 0, \quad \forall i \in I_{+0} \cup \beta_2 \\ & \nabla G_i(x^*)^T d \geq 0, \quad \forall i \in \beta_2 \\ & \nabla H_i(x^*)^T d \geq 0, \quad \forall i \in \beta_1\}. \end{aligned}$$

Logo pelo Lema 1.9

$$\begin{aligned} \mathcal{T}_{NLP_*(\beta_1, \beta_2)}(x^*)^* = \{ & v \in \mathbb{R}^n : v = - \sum_{i \in \mathcal{I}_g} u_i^g \nabla g_i(x^*) + \sum_{i=1}^p u_i^h \nabla h_i(x^*) \\ & + \sum_{i \in I_{0+} \cup I_{00}} u_i^G \nabla G_i(x^*) + \sum_{i \in I_{+0} \cup I_{00}} u_i^H \nabla H_i(x^*) \\ & u_{\mathcal{I}_g}^g \geq 0, u_{\beta_2}^G \geq 0, u_{\beta_1}^H \geq 0\}. \end{aligned} \quad (2.9)$$

Como todos os gradientes em (2.9) são linearmente independentes, temos que os escalares u^g , u^h , u^G e u^H estão univocamente determinados. Assim, considerando $v \in \mathcal{T}(x^*)^*$ arbitrário, temos por (2.8) que

$$v \in \mathcal{T}_{NLP_*(\beta_1, \beta_2)}(x^*)^* \text{ e } v \in \mathcal{T}_{NLP_*(\beta_2, \beta_1)}(x^*)^*,$$

o que implica que $u_{I_{00}}^G \geq 0$, $u_{I_{00}}^H \geq 0$. Logo:

$$\begin{aligned} v \in \{ & v \in \mathbb{R}^n : v = - \sum_{i \in \mathcal{I}_g} u_i^g \nabla g_i(x^*) + \sum_{i=1}^p u_i^h \nabla h_i(x^*) \\ & + \sum_{i \in I_{0+} \cup I_{00}} u_i^G \nabla G_i(x^*) + \sum_{i \in I_{+0} \cup I_{00}} u_i^H \nabla H_i(x^*) \\ & u_{\mathcal{I}_g}^g \geq 0, u_{I_{00}}^G \geq 0, u_{I_{00}}^H \geq 0\}. \end{aligned}$$

Finalmente, como por (1.13)

$$\begin{aligned} \mathcal{T}^{lin}(x^*) = \{ & d \in \mathbb{R}^n \mid \nabla g_i(x^*)^T d \leq 0, \quad \forall i \in \mathcal{I}_g \\ & \nabla h_i(x^*)^T d = 0, \quad \forall i = 1, \dots, p \\ & \nabla G_i(x^*)^T d = 0, \quad \forall i \in I_{0+} \\ & \nabla H_i(x^*)^T d = 0, \quad \forall i \in I_{+0} \\ & \nabla G_i(x^*)^T d \geq 0, \quad \forall i \in I_{00} \\ & \nabla H_i(x^*)^T d \geq 0, \quad \forall i \in I_{00}, \end{aligned}$$

temos pelo Lema 1.9 que

$$v \in \mathcal{T}^{lin}(x^*)^*$$

isto é $\mathcal{T}(x^*)^* \subseteq \mathcal{T}^{lin}(x^*)^*$. □

Note que da demonstração do teorema anterior, temos o seguinte resultado:

Corolário 2.13. *Se um ponto viável x^* do Problema MPEC (1) satisfaz MPEC-LICQ, então satisfaz LICQ para todo problema $NLP_*(\beta_1, \beta_2)$ e satisfaz GCQ.*

2.2.2 MPEC-ACQ e GCQ

Para ver um resultado semelhante ao caso MPEC-LICQ, agora com a condição mais fraca como MPEC-ACQ, é necessário impor uma condição ao cone viável linearizado baseada na definição a seguir.

Definição 2.14. *Dado um sistema linear*

$$Ax \leq b, \quad Cx = d. \quad (2.10)$$

Se a_i é uma linha da matriz A , uma desigualdade $a_i x \leq b_i$ se diz não singular no sistema (2.10) se existe uma solução do sistema (2.10) que satisfaz a desigualdade estrita.

Lembrando que neste caso, por (1.13),

$$\begin{aligned} \mathcal{T}^{lin}(x^*) = \{ & d \in \mathbb{R}^n \mid \nabla g_i(x^*)^T d \leq 0, \quad \forall i \in \mathcal{I}_g \\ & \nabla h_i(x^*)^T d = 0, \quad \forall i = 1, \dots, p \\ & \nabla G_i(x^*)^T d = 0, \quad \forall i \in I_{0+} \\ & \nabla H_i(x^*)^T d = 0, \quad \forall i \in I_{+0} \\ & \nabla G_i(x^*)^T d \geq 0, \quad \forall i \in I_{00} \\ & \nabla H_i(x^*)^T d \geq 0, \quad \forall i \in I_{00} \}. \end{aligned} \quad (2.11)$$

Consideremos

$$\begin{aligned} I_{00}^G &= \{i \in I_{00} : \nabla G_i(x^*)^T d \geq 0 \text{ é não singular no Sistema (2.11)}\} \\ I_{00}^H &= \{i \in I_{00} : \nabla H_i(x^*)^T d \geq 0 \text{ é não singular no Sistema (2.11)}\} \end{aligned}$$

logo

$$\begin{aligned}
\mathcal{T}^{lin}(x^*) &= \{d \in \mathbb{R}^n \mid \nabla g_i(x^*)^T d \leq 0, \forall i \in \mathcal{I}_g \\
&\quad \nabla h_i(x^*)^T d = 0, \forall i = 1, \dots, p \\
&\quad \nabla G_i(x^*)^T d = 0, \forall i \in I_{0+} \cup (I_{00} \setminus I_{00}^G) \\
&\quad \nabla H_i(x^*)^T d = 0, \forall i \in I_{+0} \cup (I_{00} \setminus I_{00}^H) \\
&\quad \nabla G_i(x^*)^T d \geq 0, \forall i \in I_{00}^G \\
&\quad \nabla H_i(x^*)^T d \geq 0, \forall i \in I_{00}^H\}.
\end{aligned} \tag{2.12}$$

Condição 2.15. [58] Dado um ponto viável x^* , existe uma partição $(\beta_1^{GH}, \beta_2^{GH}) \in P(I_{00}^G \cap I_{00}^H)$ tal que:

- Para cada $i_0 \in \beta_1^{GH}$ existe um vetor d tal que

$$\begin{aligned}
\nabla G_{i_0}(x^*)^T d &> 0 \\
\nabla G_i(x^*)^T d &= 0, \forall i \in I_{0+} \cup (I_{00} \setminus \{i_0\}) \\
\nabla H_i(x^*)^T d &= 0, \forall i \in I_{+0} \cup I_{00} \\
\nabla g_i(x^*)^T d &= 0, \forall i \in \mathcal{I}_g \\
\nabla h_i(x^*)^T d &= 0, \forall i = 1, \dots, p.
\end{aligned} \tag{2.13}$$

- Para cada $i_0 \in \beta_2^{GH}$ existe um vetor d tal que

$$\begin{aligned}
\nabla H_{i_0}(x^*)^T d &> 0 \\
\nabla H_i(x^*)^T d &= 0, \forall i \in I_{+0} \cup (I_{00} \setminus \{i_0\}) \\
\nabla G_i(x^*)^T d &= 0, \forall i \in I_{0+} \cup I_{00} \\
\nabla g_i(x^*)^T d &= 0, \forall i \in \mathcal{I}_g \\
\nabla h_i(x^*)^T d &= 0, \forall i = 1, \dots, p.
\end{aligned} \tag{2.14}$$

Teorema 2.16. [23, Teorema 4.8] Se um ponto viável x^* do Problema MPEC (1) satisfaz MPEC-ACQ e a Condição 2.15, então satisfaz GCQ.

Demonstração. Temos pelos Lemas 1.7 (b) e 1.8 (e) que $\mathcal{T}^{lin}(x^*)^* \subseteq \mathcal{T}(x^*)^*$, logo somente basta mostrar que $\mathcal{T}(x^*)^* \subseteq \mathcal{T}^{lin}(x^*)^*$. Como vale MPEC-ACQ, temos pelo Lema 2.3 que

$$\mathcal{T}(x^*) = \mathcal{T}_{MPEC}^{lin}(x^*) = \bigcup_{(\beta_1, \beta_2) \in \mathcal{P}(I_{00})} \mathcal{T}_{NLP^*(\beta_1, \beta_2)}^{lin}(x^*).$$

Logo pelo Lema 1.8 (d)

$$\mathcal{T}(x^*)^* = \mathcal{T}_{MPEC}^{lin}(x^*)^* = \bigcap_{(\beta_1, \beta_2) \in \mathcal{P}(I_{00})} \mathcal{T}_{NLP^*(\beta_1, \beta_2)}^{lin}(x^*)^*$$

assim se $v \in \mathcal{T}(x^*)^*$ então

$$v \in \mathcal{T}_{NLP_*(\beta_1, \beta_2)}^{lin}(x^*)^* \quad \forall (\beta_1, \beta_2) \in \mathcal{P}(I_{00}).$$

Em particular considerando

$$\begin{aligned} \hat{\beta}_1 &= I_{00}^H \setminus \beta_2^{GH} \text{ e} \\ \hat{\beta}_2 &= I_{00} \setminus \hat{\beta}_1 \end{aligned}$$

com $(\beta_1^{GH}, \beta_2^{GH}) \in P(I_{00}^G \cap I_{00}^H)$ satisfazendo (2.13) e (2.14), temos que $(\hat{\beta}_1, \hat{\beta}_2) \in \mathcal{P}(I_{00})$.

Logo

$$v \in \mathcal{T}_{NLP_*(\hat{\beta}_1, \hat{\beta}_2)}^{lin}(x^*)^*,$$

e como

$$\begin{aligned} \mathcal{T}_{NLP_*(\beta_1, \beta_2)}^{lin}(x^*) &= \{d \in \mathbb{R}^n \mid \nabla g_i(x^*)^T d \leq 0, \quad \forall i \in \mathcal{I}_g \\ &\quad \nabla h_i(x^*)^T d = 0, \quad \forall i = 1, \dots, p \\ &\quad \nabla G_i(x^*)^T d = 0, \quad \forall i \in I_{0+} \cup \beta_1 \\ &\quad \nabla H_i(x^*)^T d = 0, \quad \forall i \in I_{+0} \cup \beta_2 \\ &\quad \nabla G_i(x^*)^T d \geq 0, \quad \forall i \in \beta_2 \\ &\quad \nabla H_i(x^*)^T d \geq 0, \quad \forall i \in \beta_1\}, \end{aligned}$$

temos pelo Lema 1.9 que existe $u = (u^g, u^h, u^G, u^H)$ tal que

$$\begin{aligned} u_i^g &\geq 0, \quad \forall i \in \mathcal{I}_g \\ u_i^G &\geq 0, \quad \forall i \in \hat{\beta}_2 \\ u_i^H &\geq 0, \quad \forall i \in \hat{\beta}_1 \end{aligned}$$

e

$$v = - \sum_{i \in \mathcal{I}_g} u_i^g \nabla g_i(x^*) + \sum_{i=1}^p u_i^h \nabla h_i(x^*) + \sum_{i \in I_{0+} \cup I_{00}} u_i^G \nabla G_i(x^*) + \sum_{i \in I_{+0} \cup I_{00}} u_i^H \nabla H_i(x^*). \quad (2.15)$$

Agora, como $(\beta_1^{GH}, \beta_2^{GH}) \in P(I_{00}^G \cap I_{00}^H)$, então

$$\begin{aligned} I_{00}^H \setminus \beta_2^{GH} &= \hat{\beta}_1 \\ I_{00}^G \setminus \beta_1^{GH} &= I_{00}^G \setminus \hat{\beta}_1 \subseteq I_{00} \setminus \hat{\beta}_1 = \hat{\beta}_2. \end{aligned}$$

As partições $(\hat{\beta}_1, \hat{\beta}_2)$ de I_{00} , os conjuntos I_{00}^G e I_{00}^H e a partição $(\beta_1^{GH}, \beta_2^{GH})$ de $I_{00}^G \cap I_{00}^H$, podem ser vistas na Figura 2.3.

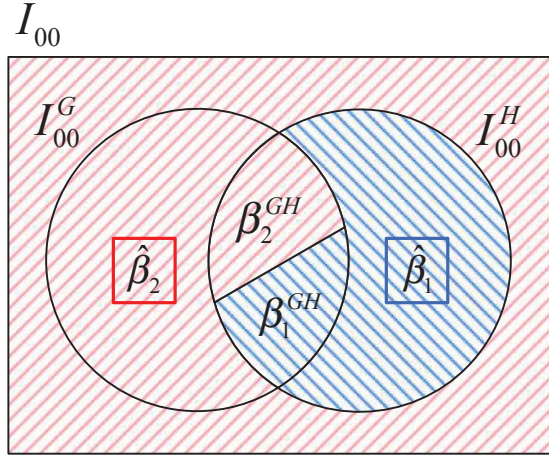


Figura 2.3: Partição $(\hat{\beta}_1, \hat{\beta}_2)$ de I_{00} , conjuntos I_{00}^G e I_{00}^H e partição $(\beta_1^{GH}, \beta_2^{GH})$ de $I_{00}^G \cap I_{00}^H$.

Logo em (2.15)

$$\begin{aligned} u_i^G &\geq 0, \quad \forall i \in I_{00}^G \setminus \beta_1^{GH} \\ u_i^H &\geq 0, \quad \forall i \in I_{00}^H \setminus \beta_2^{GH}. \end{aligned}$$

Considerando o Lema 1.9 aplicado a (2.12), temos que resta mostrar que $u_i^G \geq 0, \forall i \in \beta_1^{GH}$ e $u_i^H \geq 0, \forall i \in \beta_2^{GH}$. Para isso, sem perda de generalidade, suponhamos por contradição que $u_{i_0}^G < 0$, então por (2.13) existe d tal que

$$\begin{aligned} \nabla G_{i_0}(x^*)^T d &> 0 \\ \nabla G_i(x^*)^T d &= 0, \quad \forall i \in I_{0+} \cup (I_{00} \setminus \{i_0\}) \\ \nabla H_i(x^*)^T d &= 0, \quad \forall i \in I_{+0} \cup I_{00} \\ \nabla g_i(x^*)^T d &= 0, \quad \forall i \in \mathcal{I}_g \\ \nabla h_i(x^*)^T d &= 0, \quad \forall i = 1, \dots, p. \end{aligned}$$

Logo

$$\begin{aligned} v^T d &= \left(- \sum_{i \in \mathcal{I}_g} u_i^g \nabla g_i(x^*) + \sum_{i=1}^p u_i^h \nabla h_i(x^*) \right. \\ &\quad \left. + \sum_{i \in I_{0+} \cup I_{00}} u_i^G \nabla G_i(x^*) + \sum_{i \in I_{+0} \cup I_{00}} u_i^H \nabla H_i(x^*) \right)^T d \\ &= u_{i_0}^G \nabla G_{i_0}(x^*)^T d < 0 \end{aligned} \tag{2.16}$$

e como

$$\begin{aligned}\mathcal{T}_{NLP^*(\beta_1, \beta_2)}^{lin}(x^*) &= \{d \in \mathbb{R}^n \mid \nabla g_i(x^*)^T d \leq 0, \forall i \in \mathcal{I}_g \\ &\quad \nabla h_i(x^*)^T d = 0, \forall i = 1, \dots, p \\ &\quad \nabla G_i(x^*)^T d = 0, \forall i \in I_{0+} \cup \beta_1 \\ &\quad \nabla H_i(x^*)^T d = 0, \forall i \in I_{+0} \cup \beta_2 \\ &\quad \nabla G_i(x^*)^T d \geq 0, \forall i \in \beta_2 \\ &\quad \nabla H_i(x^*)^T d \geq 0, \forall i \in \beta_1\},\end{aligned}$$

temos que $d \in \mathcal{T}_{NLP^*(\beta_1, \beta_2)}^{lin}(x^*)$ para toda partição $(\beta_1, \beta_2) \in \mathcal{P}(I_{00})$ tal que $i_0 \in \beta_2$. Assim para toda partição desse tipo temos que $v \in \mathcal{T}_{NLP^*(\beta_1, \beta_2)}^{lin}(x^*)^*$, de onde $v^T d \geq 0$, o que contradiz (2.16). \square

A respeito da Condição 2.15, temos pelo Lema 1 de [58] que é equivalente a que se

$$\sum_{i \in \mathcal{I}_g} u_i^g \nabla g_i(x^*) + \sum_{i=1}^p u_i^h \nabla h_i(x^*) + \sum_{i \in I_{0+} \cup I_{00}} u_i^G \nabla G_i(x^*) + \sum_{i \in I_{+0} \cup I_{00}} u_i^H \nabla H_i(x^*) = 0$$

então

$$\begin{aligned}u_i^G &= 0 \quad \forall i \in \beta_1^{GH} \\ u_i^H &= 0 \quad \forall i \in \beta_2^{GH},\end{aligned}$$

o que implica a necessidade de o conjunto $\{\nabla G_i(x^*), \nabla H_j(x^*) \mid i \in \beta_1^{GH} \text{ e } j \in \beta_2^{GH}\}$ ser linearmente independente, o que é razoavelmente aceitável. Por exemplo, a condição se satisfaz para o minimizador x^* do Exemplo 1.22, pois neste caso $I_{0+} = I_{+0} = \emptyset$, $I_{00} = I_{00}^H = I_{00}^G = \{1\}$, $\nabla G_1(x^*) = (1 \ 0)^T$ e $\nabla H_1(x^*) = (0 \ 1)^T$. Porém, esta condição não é necessária, como podemos ver no seguinte exemplo:

Exemplo 2.17. *Considere o problema*

$$\begin{aligned}\min \quad & f(x) = x_1^2 + x_2^2 + x_3^2 \\ \text{s.a} \quad & h_1(x) = x_3 = 0 \\ & h_2(x) = x_2 = 0 \\ & G(x) = x_2 + x_3^2 \geq 0 \\ & H(x) = x_2 - x_3^2 \geq 0 \\ & \Theta(x) = G(x)H(x) = (x_2 + x_3^2)(x_2 - x_3^2) = 0.\end{aligned}$$

O conjunto viável é $\left\{ \begin{pmatrix} t & 0 & 0 \end{pmatrix}^T \mid t \in \mathbb{R} \right\}$ e $x^* = \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}^T$ é seu único minimi-

zador. Claramente temos $\mathcal{T}(x^*) = \left\{ \begin{pmatrix} t & 0 & 0 \end{pmatrix}^T \mid t \in \mathbb{R} \right\}$. Além disso, como $\nabla h_1(x^*) = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}^T$, $\nabla h_2(x^*) = \nabla G(x^*) = \nabla H(x^*) = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}^T$ e $\nabla \Theta(x^*) = \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}^T$, então $\mathcal{T}^{lin}(x^*) = \left\{ \begin{pmatrix} t & 0 & 0 \end{pmatrix}^T \mid t \in \mathbb{R} \right\}$. Logo, x^* satisfaz ACQ. Porém, a Condição 2.15 não se satisfaz no minimizador x^* , pois $\nabla G_i(x^*) = \nabla H_i(x^*)$ para todo $i \in I_{00} = I_{00}^G = I_{00}^H = \{1\}$.

2.3 Estacionariedade em MPEC

Como veremos posteriormente, KKT é uma condição de estacionariedade se considerarmos MPEC-LICQ, porém não é uma condição necessária de otimalidade com as CQs mais fracas MPEC-MFCQ, MPEC-ACQ e MPEC-GCQ. Nesse sentido, em contraste com a programação não linear usual, surgiram diferentes definições de estacionariedade para problemas MPEC. No que segue, nos restringiremos aqueles importantes para nossa análise, descrevendo a relação existente entre esses conceitos, bem como suas relação com as CQs MPEC da seção anterior.

Definição 2.18. *Seja x^* um ponto viável do Problema MPEC (1). O ponto x^* se diz:*

1. *Fracamente estacionário (W -estacionário) se existem multiplicadores $\lambda^* = (\lambda^g, \lambda^h, \lambda^G, \lambda^H) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^l \times \mathbb{R}^l$ tais que*

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i^g \nabla g_i(x^*) + \sum_{i=1}^p \lambda_i^h \nabla h_i(x^*) - \sum_{i=1}^l \lambda_i^G \nabla G_i(x^*) - \sum_{i=1}^l \lambda_i^H \nabla H_i(x^*) = 0$$

com

$$\begin{aligned} \lambda_{I_{0+}}^G &= 0, & \lambda_{I_{0+}}^H &= 0, \\ \lambda^g &\geq 0, & g(x^*) &\leq 0, & (\lambda^g)^T g(x^*) &= 0; \end{aligned}$$

2. *Clarke-estacionário (C -estacionário) se é fracamente estacionário e $\lambda_i^G \lambda_i^H \geq 0$ para todo $i \in I_{00}$;*
3. *Mordukhovich-estacionário (M -estacionário) se é fracamente estacionário e $\lambda_i^G > 0$, $\lambda_i^H > 0$ ou $\lambda_i^G \lambda_i^H = 0$ para todo $i \in I_{00}$;*
4. *Fortemente-estacionário (S -estacionário) se é fracamente estacionário e $\lambda_i^G \geq 0$, $\lambda_i^H \geq 0$ para todo $i \in I_{00}$ (KKT para MPEC).*

As diferentes condições dos multiplicadores λ_i^G e λ_i^H com $i \in I_{00}$ estão ilustradas na Figura 2.4. Note que:

$$\text{S-est.} \implies \text{M-est.} \implies \text{C-est.} \implies \text{Fracamente est.}$$

O seguinte resultado mostra que os conceitos KKT e de S-estacionariedade são equivalentes.

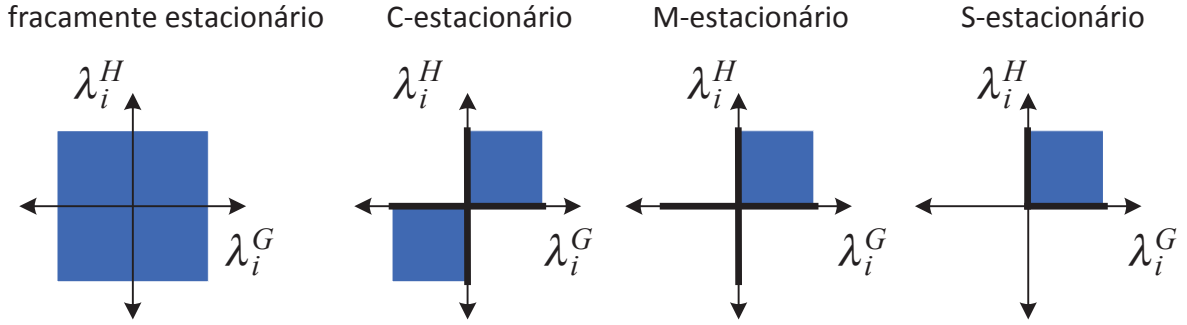


Figura 2.4: Ilustração geométrica dos valores que os multiplicadores de Lagrange podem assumir para os diferentes conceitos de estacionariedade.

Proposição 2.19. [23, Proposição 4.2] *Seja x^* um ponto viável do Problema MPEC (1). Então existe um multiplicador de Lagrange $\hat{\lambda}$ tal que $(x^*, \hat{\lambda})$ satisfaz as condições KKT para o Problema MPEC (1), se e somente se, existe um multiplicador de Lagrange λ^* tal que (x^*, λ^*) satisfaz a condição de estacionariedade forte.*

Demonstração. Considerando $\theta(x) = G(x)^T H(x)$, temos que as condições de KKT para o problema

$$\begin{aligned} \min \quad & f(x) \\ \text{s.a} \quad & g(x) \leq 0, \quad h(x) = 0 \\ & G(x) \geq 0, \quad H(x) \geq 0, \quad G(x)^T H(x) = 0 \end{aligned}$$

implicam

$$\nabla f(x^*) + \sum_{i=1}^m \hat{\lambda}_i^g \nabla g_i(x^*) + \sum_{i=1}^p \hat{\lambda}_i^h \nabla h_i(x^*) - \sum_{i=1}^l \hat{\lambda}_i^G \nabla G_i(x^*) - \sum_{i=1}^l \hat{\lambda}_i^H \nabla H_i(x^*) + \hat{\lambda}^\theta \nabla \theta(x^*) = 0$$

com

$$\begin{aligned} G(x^*) \geq 0 \quad \hat{\lambda}^G \geq 0 \quad & (\hat{\lambda}^G)^T G(x^*) = 0 \\ H(x^*) \geq 0 \quad \hat{\lambda}^H \geq 0 \quad & (\hat{\lambda}^H)^T H(x^*) = 0 \\ g(x^*) \leq 0 \quad \hat{\lambda}^g \geq 0 \quad & (\hat{\lambda}^g)^T g(x^*) = 0. \end{aligned}$$

Como

$$\nabla \theta(x^*) = \sum_{i=1}^l (G_i(x^*) \nabla H_i(x^*) + H_i(x^*) \nabla G_i(x^*)),$$

temos que

$$\begin{aligned} \nabla f(x^*) + \sum_{i=1}^m \hat{\lambda}_i^g \nabla g_i(x^*) + \sum_{i=1}^p \hat{\lambda}_i^h \nabla h_i(x^*) - \sum_{i=1}^l \hat{\lambda}_i^G \nabla G_i(x^*) - \sum_{i=1}^l \hat{\lambda}_i^H \nabla H_i(x^*) + \\ \hat{\lambda}_i^\theta \sum_{i=1}^l (G_i(x^*) \nabla H_i(x^*) + H_i(x^*) \nabla G_i(x^*)) = 0, \end{aligned}$$

de onde

$$\begin{aligned} \nabla f(x^*) + \sum_{i=1}^m \hat{\lambda}_i^g \nabla g_i(x^*) + \sum_{i=1}^p \hat{\lambda}_i^h \nabla h_i(x^*) - \sum_{i=1}^l (\hat{\lambda}_i^G - \hat{\lambda}^\theta H_i(x^*)) \nabla G_i(x^*) - \\ \sum_{i=1}^l (\hat{\lambda}_i^H - \hat{\lambda}^\theta G_i(x^*)) \nabla H_i(x^*) = 0. \end{aligned}$$

Logo considerando

$$\lambda^g = \hat{\lambda}^g, \quad \lambda^h = \hat{\lambda}^h, \quad \lambda_i^G = \hat{\lambda}_i^G - \hat{\lambda}^\theta H_i(x^*), \quad \lambda_i^H = \hat{\lambda}_i^H - \hat{\lambda}^\theta G_i(x^*)$$

temos que o par $(x^*, (\lambda^g, \lambda^h, \lambda^G, \lambda^H))$ satisfaz as condições de estacionariedade forte, pois

$$\begin{aligned} \lambda_{I_{00}}^G &= \hat{\lambda}_{I_{00}}^G - \hat{\lambda}^\theta H_{I_{00}}(x^*) = \hat{\lambda}_{I_{00}}^G \geq 0 \\ \lambda_{I_{00}}^H &= \hat{\lambda}_{I_{00}}^H - \hat{\lambda}^\theta G_{I_{00}}(x^*) = \hat{\lambda}_{I_{00}}^H \geq 0 \\ \lambda_{I_{+0}}^G &= \hat{\lambda}_{I_{+0}}^G - \hat{\lambda}^\theta H_{I_{+0}}(x^*) = \hat{\lambda}_{I_{+0}}^G = 0 \\ \lambda_{I_{0+}}^H &= \hat{\lambda}_{I_{0+}}^H - \hat{\lambda}^\theta G_{I_{0+}}(x^*) = \hat{\lambda}_{I_{0+}}^H = 0. \end{aligned}$$

Para a recíproca considere um par S-estacionário (x^*, λ^*) com $\lambda^* = (\lambda^g, \lambda^h, \lambda^G, \lambda^H)$, isto é

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i^g \nabla g_i(x^*) + \sum_{i=1}^p \lambda_i^h \nabla h_i(x^*) - \sum_{i=1}^l \lambda_i^G \nabla G_i(x^*) - \sum_{i=1}^l \lambda_i^H \nabla H_i(x^*) = 0 \quad (2.17)$$

com

$$\begin{aligned} \lambda_{I_{+0}}^G = 0 \quad \lambda_{I_{0+}}^H = 0 \quad \lambda_i^G \geq 0, \lambda_i^H \geq 0 \text{ para todo } i \in I_{00} \\ \lambda^g \geq 0 \quad g(x^*) \leq 0 \quad (\lambda^g)^T g(x^*) = 0. \end{aligned}$$

Logo considerando

$$\hat{\lambda}^g = \lambda^g, \quad \hat{\lambda}^h = \lambda^h, \quad \hat{\lambda}_{I_{00} \cup I_{+0}}^G = \lambda_{I_{00} \cup I_{+0}}^G, \quad \hat{\lambda}_{I_{00} \cup I_{0+}}^H = \lambda_{I_{00} \cup I_{0+}}^H \quad (2.18)$$

e

$$\hat{\lambda}^\theta = \max \left\{ \max_{i \in I_{0+}} \left\{ \frac{-\lambda_i^G}{H_i(x^*)} \right\}, \max_{i \in I_{+0}} \left\{ \frac{-\lambda_i^H}{G_i(x^*)} \right\} \right\}$$

temos que

$$\hat{\lambda}^\theta + \frac{\lambda_i^G}{H_i(x^*)} \geq 0 \quad \text{e} \quad \hat{\lambda}^\theta + \frac{\lambda_j^H}{G_j(x^*)} \geq 0$$

para todos $i \in I_{0+}$ e $j \in I_{+0}$.

Isto permite definir

$$\hat{\lambda}_{I_{0+}}^G = \lambda_{I_{0+}}^G + \hat{\lambda}^\theta H_{I_{0+}}(x^*) \geq 0 \quad \text{e} \quad \hat{\lambda}_{I_{+0}}^H = \lambda_{I_{+0}}^H + \hat{\lambda}^\theta G_{I_{+0}}(x^*) \geq 0. \quad (2.19)$$

Desta forma, como

$$\begin{aligned}\nabla\theta(x^*) &= \sum_{i=1}^l (G_i(x^*)\nabla H_i(x^*) + H_i(x^*)\nabla G_i(x^*)) \\ &= \sum_{i \in I_{+0}} G_i(x^*)\nabla H_i(x^*) + \sum_{i \in I_{0+}} H_i(x^*)\nabla G_i(x^*),\end{aligned}$$

temos de (2.18), (2.19) e (2.17) que

$$\nabla f(x^*) + \sum_{i=1}^m \hat{\lambda}_i^g \nabla g_i(x^*) + \sum_{i=1}^p \hat{\lambda}_i^h \nabla h_i(x^*) - \sum_{i=1}^l \hat{\lambda}_i^G \nabla G_i(x^*) - \sum_{i=1}^l \hat{\lambda}_i^H \nabla H_i(x^*) + \hat{\lambda}^\theta \nabla\theta(x^*) = 0$$

resultando que $(x^*, (\hat{\lambda}^g, \hat{\lambda}^h, \hat{\lambda}^G, \hat{\lambda}^H, \hat{\lambda}^\theta))$ satisfaz as condições *KKT*. \square

2.4 Condições de qualificação para MPEC e estacionariedade

Como a CQ de Guignard garante que um minimizador satisfaz *KKT* [60] então pela Proposição 2.19 também garante que o minimizador é *S*-estacionário. Mas o que se pode dizer das condições de qualificação para MPEC e seu relacionamento com os conceitos de estacionariedade? Apresentamos alguns resultados importantes que relacionam estes conceitos.

2.4.1 MPEC-LICQ e estacionariedade

Nas seções anteriores provamos no Teorema 2.12 que MPEC-LICQ implica GCQ, e na Proposição 2.19 que *KKT* é equivalente a *S*-estacionário, logo MPEC-LICQ é uma condição de primeira ordem para que um minimizador do Problema MPEC (1) seja *S*-estacionário. Porém veremos uma prova alternativa para familiarizarmos com esses novos conceitos de estacionariedade.

Teorema 2.20. [24, Teorema 3.1] *Seja x^* um minimizador local do MPEC (1). Se x^* satisfaz MPEC-LICQ então x^* é *S*-estacionário.*

Demonstração. Consideremos uma partição $(\beta_1, \beta_2) \in I_{00}$ e os Problemas $NLP_*(\beta_1, \beta_2)$ e $NLP_*(\beta_2, \beta_1)$ (1.14). O ponto x^* é um minimizador destes dois problemas pois é viável para ambos e seus conjuntos viáveis são subconjuntos do conjunto viável associado ao Problema original (1). Logo como vale MPEC-LICQ em x^* , temos do Corolário 2.13 que x^* satisfaz LICQ para todo problema $NLP_*(\beta_1, \beta_2)$. Em consequência vale *KKT*,

existindo multiplicadores $\tilde{\lambda}$ e $\hat{\lambda}$ tais que

$$\begin{aligned} \nabla f(x^*) + \sum_{i=1}^m \tilde{\lambda}_i^g \nabla g_i(x^*) + \sum_{i=1}^p \tilde{\lambda}_i^h \nabla h_i(x^*) - \sum_{i=1}^l \left[\tilde{\lambda}_i^G \nabla G_i(x^*) + \tilde{\lambda}_i^H \nabla H_i(x^*) \right] &= 0 \\ G_{I_{0+} \cup \beta_1}(x^*) = 0 \quad G_{I_{+0} \cup \beta_2}(x^*) &\geq 0 \quad \tilde{\lambda}_{I_{+0} \cup \beta_2}^G \geq 0 \\ H_{I_{+0} \cup \beta_2}(x^*) = 0 \quad H_{I_{0+} \cup \beta_1}(x^*) &\geq 0 \quad \tilde{\lambda}_{I_{0+} \cup \beta_1}^H \geq 0 \\ h(x^*) = 0 \quad g(x^*) &\leq 0 \quad \tilde{\lambda}^g \geq 0 \\ (\tilde{\lambda}_{I_{+0} \cup \beta_2}^G)^T G_{I_{+0} \cup \beta_2}(x^*) = 0 \quad (\tilde{\lambda}_{I_{0+} \cup \beta_1}^H)^T H_{I_{0+} \cup \beta_1}(x^*) &= 0 \quad (\tilde{\lambda}^g)^T g(x^*) = 0 \end{aligned} \quad (2.20)$$

e

$$\begin{aligned} \nabla f(x^*) + \sum_{i=1}^m \hat{\lambda}_i^g \nabla g_i(x^*) + \sum_{i=1}^p \hat{\lambda}_i^h \nabla h_i(x^*) - \sum_{i=1}^l \left[\hat{\lambda}_i^G \nabla G_i(x^*) + \hat{\lambda}_i^H \nabla H_i(x^*) \right] &= 0 \\ G_{I_{0+} \cup \beta_2}(x^*) = 0 \quad G_{I_{+0} \cup \beta_1}(x^*) &\geq 0 \quad \hat{\lambda}_{I_{+0} \cup \beta_1}^G \geq 0 \\ H_{I_{+0} \cup \beta_2}(x^*) = 0 \quad H_{I_{0+} \cup \beta_1}(x^*) &\geq 0 \quad \hat{\lambda}_{I_{0+} \cup \beta_2}^H \geq 0 \\ h(x^*) = 0 \quad g(x^*) &\leq 0 \quad \hat{\lambda}^g \geq 0 \\ (\hat{\lambda}_{I_{+0} \cup \beta_1}^G)^T G_{I_{+0} \cup \beta_1}(x^*) = 0 \quad (\hat{\lambda}_{I_{0+} \cup \beta_2}^H)^T H_{I_{0+} \cup \beta_2}(x^*) &= 0 \quad (\hat{\lambda}^g)^T g(x^*) = 0. \end{aligned} \quad (2.21)$$

Agora como $G_{I_{+0}}(x^*) > 0$ e $H_{I_{0+}}(x^*) > 0$, pela definição de I_{+0} e I_{0+} , temos pelas equações de complementaridade em (2.20) e (2.21) que

$$\tilde{\lambda}_{I_{+0}}^G = \hat{\lambda}_{I_{+0}}^G = 0 \quad \text{e} \quad \tilde{\lambda}_{I_{0+}}^H = \hat{\lambda}_{I_{0+}}^H = 0$$

e similarmente $\tilde{\lambda}_i^g = \hat{\lambda}_i^g = 0$ para todo $i \in \mathcal{I}_g$, logo igualando os gradientes das equações anteriores temos que

$$\begin{aligned} 0 &= \sum_{i=1}^m \left(\tilde{\lambda}_i^g - \hat{\lambda}_i^g \right) \nabla g_i(x^*) + \sum_{i=1}^p \left(\tilde{\lambda}_i^h - \hat{\lambda}_i^h \right) \nabla h_i(x^*) \\ &\quad - \sum_{i=1}^l \left(\tilde{\lambda}_i^G - \hat{\lambda}_i^G \right) \nabla G_i(x^*) - \sum_{i=1}^l \left(\tilde{\lambda}_i^H - \hat{\lambda}_i^H \right) \nabla H_i(x^*) \\ &= \sum_{i \in \mathcal{I}_g} \left(\tilde{\lambda}_i^g - \hat{\lambda}_i^g \right) \nabla g_i(x^*) + \sum_{i=1}^p \left(\tilde{\lambda}_i^h - \hat{\lambda}_i^h \right) \nabla h_i(x^*) \\ &\quad - \sum_{i \in I_{0+} \cup I_{00}} \left(\tilde{\lambda}_i^G - \hat{\lambda}_i^G \right) \nabla G_i(x^*) - \sum_{i \in I_{+0} \cup I_{00}} \left(\tilde{\lambda}_i^H - \hat{\lambda}_i^H \right) \nabla H_i(x^*) \end{aligned}$$

e como vale MPEC-LICQ em x^* , todos os gradientes são linearmente independentes, resultando que $\tilde{\lambda} = \hat{\lambda}$, podendo-se assim definir $\lambda^* = \tilde{\lambda} = \hat{\lambda}$ que possui as duas propriedades dos multiplicadores de (2.20) e (2.21), resultando que:

$$\begin{aligned} (\lambda_{I_{00}}^G)^* &\geq 0 \quad (\lambda_{I_{+0}}^G)^* = 0 \\ (\lambda_{I_{00}}^H)^* &\geq 0 \quad (\lambda_{I_{0+}}^H)^* = 0 \end{aligned}$$

o que conclui que x^* é S-estacionário. \square

2.4.2 MPEC-MFCQ e estacionariedade

Na seção anterior mostramos que, se x^* é um minimizador do Problema MPEC (1), MPEC-LICQ implica S-estacionariedade para x^* . Porém MPEC-MFCQ não pode garantir o mesmo, como pode ser visto no seguinte exemplo.

Exemplo 2.21. [65] Considere o seguinte problema MPEC

$$\begin{aligned} \min \quad & x_1 + x_2 - x_3 \\ \text{s.a} \quad & g_1(x) = -4x_1 + x_3 \leq 0 \\ & g_2(x) = -4x_2 + x_3 \leq 0 \\ & G(x) = x_1 \geq 0 \\ & H(x) = x_2 \geq 0 \\ & G(x)^T H(x) = x_1 x_2 = 0. \end{aligned}$$

Como $x_1 \geq 0$ e $x_2 \geq 0$ então o mínimo acontece quando $x_1 = x_2 = 0$, resultando que $x_3 = 0$. Assim $x^* = \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}^T$ é o único minimizador do problema. Temos que $I_{0+} = \emptyset$, $I_{+0} = \emptyset$ e $I_{00} = \{1\}$, $\nabla G(x^*) = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}^T$ e $\nabla H(x^*) = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}^T$ são linearmente independentes e que existe $\begin{pmatrix} 0 & 0 & -1 \end{pmatrix}^T$ tal que

$$\begin{aligned} \nabla g_1(x^*)^T d &= \begin{pmatrix} -4 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & -1 \end{pmatrix}^T < 0, \\ \nabla g_2(x^*)^T d &= \begin{pmatrix} 0 & -4 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & -1 \end{pmatrix}^T < 0, \\ \nabla G(x^*)^T d &= \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & -1 \end{pmatrix}^T = 0, \\ \nabla H(x^*)^T d &= \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & -1 \end{pmatrix}^T = 0, \end{aligned}$$

logo x^* satisfaz MPEC-MFCQ. Por outro lado, se

$$\nabla f(x^*) + \lambda_1^g \nabla g_1(x^*) + \lambda_2^g \nabla g_2(x^*) - \lambda^G \nabla G(x^*) - \lambda^H \nabla H(x^*) = 0$$

então

$$\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} + \lambda_1^g \begin{pmatrix} -4 \\ 0 \\ 1 \end{pmatrix} + \lambda_2^g \begin{pmatrix} 0 \\ -4 \\ 1 \end{pmatrix} - \lambda^G \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \lambda^H \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = 0,$$

que equivale a

$$\begin{cases} 1 - 4\lambda_1^g - \lambda^G &= 0 \\ 1 - 4\lambda_2^g - \lambda^H &= 0 \\ -1 + \lambda_1^g + \lambda_2^g &= 0 \end{cases}$$

resultando que

$$\lambda^G + \lambda^H = -2.$$

Assim x^* não é S-estacionário. Porém se $(\lambda^G, \lambda^H) \in \{(0, -2), (-2, 0)\}$ então a condição de M-estacionariedade é satisfeita posto que neste caso $\lambda^G \lambda^H = 0$. Além disso se considerarmos os multiplicadores no segmento de reta $\{(\lambda^G, \lambda^H) \mid \lambda^G \in [-2, 0] \text{ e } \lambda^G + \lambda^H = -2\}$, então se satisfaz a condição de C-estacionariedade ($\lambda^G \lambda^H \geq 0$). Destacamos também deste exemplo que nem mesmo as restrições lineares (exceto a complementaridade) garantem que todo minimizador é S-estacionário. A Figura 2.5 mostra os multiplicadores de Lagrange do Exemplo 2.21.

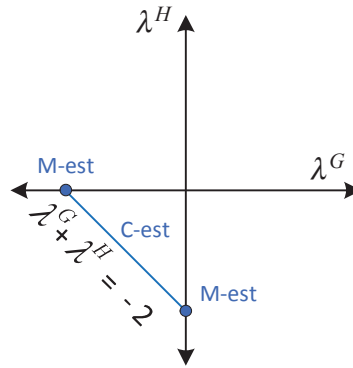


Figura 2.5: Multiplicadores de Lagrange do Exemplo 2.21. Todos os pontos do segmento são C-estacionários e os extremos são M-estacionários.

2.4.3 MPEC-GCQ e estacionariedade

O Exemplo 2.21 sugere a possibilidade de que quando o minimizador satisfaz MPEC-MFCQ, então o ponto é M-estacionário. No que segue mostraremos que este resultado vale inclusive considerando uma condição mais fraca como MPEC-GCQ. Para isso seguiremos a demonstração feita em [22], sendo necessário introduzir alguns conceitos e resultados prévios.

Definição 2.22. [22, Definição 2.2] Seja $\Omega \subseteq \mathbb{R}^l$ um conjunto não vazio e fechado. Se $v \in \Omega$, definimos o cone normal de Fréchet de Ω em v como o conjunto

$$\hat{N}(v, \Omega) = \left\{ w \in \mathbb{R}^l \mid \limsup_{\substack{v^k \rightarrow v \\ \{v^k\} \subseteq \Omega \setminus \{v\}}} \frac{w^T(v^k - v)}{\|v^k - v\|} \leq 0 \right\}.$$

e o cone normal limite de Ω em v como o conjunto

$$N(v, \Omega) = \left\{ w \mid \exists \{v^k\} \text{ em } \Omega \text{ e } \{w^k\} \text{ em } \hat{N}(v^k, \Omega) \text{ com } \lim_{k \rightarrow \infty} w^k = w, \lim_{k \rightarrow \infty} v^k = v \right\}.$$

Proposição 2.23. [64, Proposição 6.5 e Teorema 6.28] *Seja $\Omega \subseteq \mathbb{R}^n$ o conjunto viável do Problema MPEC (1). Se $\Omega \neq \emptyset$ e $x^* \in \Omega$ então*

$$-\mathcal{T}(x^*)^* = \hat{N}(x^*, \Omega) \subseteq N(x^*, \Omega).$$

Proposição 2.24. [64, Teorema 6.9] *Se Ω é convexo e $x^* \in \Omega$, então $\hat{N}(x^*, \Omega) = N(x^*, \Omega)$.*

Proposição 2.25. [64, Proposição 6.41] *Seja $\Omega_1 \subseteq \mathbb{R}^{l_1}$ e $\Omega_2 \subseteq \mathbb{R}^{l_2}$ não vazios e fechados. Então para $v_1 \in \Omega_1$ e $v_2 \in \Omega_2$ temos*

$$N((v_1, v_2), \Omega_1 \times \Omega_2) = N(v_1, \Omega_1) \times N(v_2, \Omega_2).$$

Proposição 2.26. [25, Proposição 2.4] *Se o conjunto \mathcal{C} está definido por*

$$\mathcal{C} = \{(a, b) \in \mathbb{R}^{2l} : a \geq 0, b \geq 0 \text{ e } a^T b = 0\}$$

então, considerando $(a, b) \in \mathcal{C}$ e

$$\mathcal{I}_a = \{i : a_i = 0, b_i > 0\}, \quad \mathcal{I}_b = \{i : a_i > 0, b_i = 0\}, \quad \mathcal{I}_{ab} = \{i : a_i = 0, b_i = 0\},$$

temos que o cone normal limite de \mathcal{C} está dado por

$$N((a, b), \mathcal{C}) = \{(x, y) \in \mathbb{R}^{2l} : x_{\mathcal{I}_b} = 0, y_{\mathcal{I}_a} = 0, ((x_i < 0 \text{ e } y_i < 0) \text{ ou } x_i y_i = 0 \quad \forall i \in \mathcal{I}_{ab})\}.$$

Definição 2.27. *Dados dos conjuntos X e Y , uma multifunção F de X em Y , denotada por $F : X \rightrightarrows Y$, é uma função de X nas partes de Y , com a propriedade $F(x) \neq \emptyset$ para todo $x \in X$. Sendo seu gráfico, o conjunto:*

$$\text{gph} F = \{(x, y) \in X \times Y \mid y \in F(x)\}.$$

Definição 2.28. *Dada uma matriz A de ordem $m \times n$ e um vetor $b \in \mathbb{R}^m$, um conjunto poliédrico $P \subseteq \mathbb{R}^n$ é um conjunto da forma:*

$$P = \{x \in \mathbb{R}^n \mid Ax \leq b\}.$$

Proposição 2.29. [36, Corolário 4.2] *Sejam C_1 e C_2 conjuntos arbitrários e $x \in C_1 \cap C_2$. Se o gráfico da multifunção $\Phi : \mathbb{R}^p \rightrightarrows \mathbb{R}^q$ dada por*

$$\Phi(v) = \{w \in C_1 : v + w \in C_2\}$$

é a união de um número finito de conjuntos poliédricos convexos, então

$$N(x, C_1 \cap C_2) \subseteq N(x, C_1) + N(x, C_2).$$

Como é conhecido na literatura, GCQ é a CQ mais fraca que garante que um minimizador cumpra KKT. O resultado a seguir mostra um resultado similar considerando MPEC-GCQ para minimizadores do Problema MPEC (1).

Teorema 2.30. [22, Teorema 3.1] *Seja x^* um minimizador do Problema MPEC (1). Se x^* satisfaz MPEC-GCQ, então x^* é M-estacionário, isto é, existem multiplicadores $\lambda^* = (\lambda^g, \lambda^h, \lambda^G, \lambda^H)$ tais que*

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i^g \nabla g_i(x^*) + \sum_{i=1}^p \lambda_i^h \nabla h_i(x^*) - \sum_{i=1}^l \lambda_i^G \nabla G_i(x^*) - \sum_{i=1}^l \lambda_i^H \nabla H_i(x^*) = 0$$

com

$$\begin{aligned} \lambda^g &\geq 0, \quad g(x^*) \leq 0, \quad (\lambda^g)^T g(x^*) = 0, \\ \lambda_{I_{+0}}^G &= 0, \quad \lambda_{I_{0+}}^H = 0, \quad (\lambda_i^G > 0 \text{ e } \lambda_i^H > 0) \text{ ou } \lambda_i^G \lambda_i^H = 0 \quad \forall i \in I_{00}. \end{aligned}$$

Demonstração. Se x^* é um minimizador então pelo Lema 1.7 (c)

$$\nabla f(x^*)^T d \geq 0 \quad \forall d \in \mathcal{T}(x^*),$$

logo $\nabla f(x^*) \in \mathcal{T}(x^*)^*$. Como vale MPEC-GCQ

$$\nabla f(x^*) \in \mathcal{T}_{MPEC}^{lin}(x^*)^*$$

e assim

$$\nabla f(x^*)^T d \geq 0 \quad \forall d \in \mathcal{T}_{MPEC}^{lin}(x^*),$$

o que equivale a dizer que $d^* = 0$ é um minimizador local de

$$\begin{aligned} \min_d \quad & \nabla f(x^*)^T d \\ \text{s.a} \quad & d \in \mathcal{T}_{MPEC}^{lin}(x^*). \end{aligned} \tag{2.22}$$

Lembrando que

$$\begin{aligned}
\mathcal{T}_{MPEC}^{lin}(x^*) &= \{d \in \mathbb{R}^n \mid \nabla g_i(x^*)^T d \leq 0, \forall i \in \mathcal{I}_g \\
&\quad \nabla h_i(x^*)^T d = 0, \forall i = 1, \dots, p \\
&\quad \nabla G_i(x^*)^T d = 0, \forall i \in I_{0+} \\
&\quad \nabla H_i(x^*)^T d = 0, \forall i \in I_{+0} \\
&\quad \nabla G_i(x^*)^T d \geq 0, \forall i \in I_{00} \\
&\quad \nabla H_i(x^*)^T d \geq 0, \forall i \in I_{00} \\
&\quad (\nabla G_i(x^*)^T d) (\nabla H_i(x^*)^T d) = 0, \forall i \in I_{00}\}
\end{aligned}$$

temos que (2.22) equivale a que $(d^*, \xi^*, \eta^*) = (0, 0, 0)$ seja um minimizador de

$$\begin{aligned}
&\min_{(d, \xi, \eta)} \quad \nabla f(x^*)^T d \\
&s.a \quad (d, \xi, \eta) \in \mathcal{D} = \mathcal{D}_1 \cap \mathcal{D}_2
\end{aligned} \tag{2.23}$$

onde

$$\begin{aligned}
\mathcal{D}_1 &= \{(d, \xi, \eta) \in \mathbb{R}^{n+2|I_{00}|} \mid \nabla g_i(x^*)^T d \leq 0, \forall i \in \mathcal{I}_g \\
&\quad \nabla h_i(x^*)^T d = 0, \forall i = 1, \dots, p \\
&\quad \nabla G_i(x^*)^T d = 0, \forall i \in I_{0+} \\
&\quad \nabla H_i(x^*)^T d = 0, \forall i \in I_{+0} \\
&\quad \nabla G_i(x^*)^T d - \xi_i = 0, \forall i \in I_{00} \\
&\quad \nabla H_i(x^*)^T d - \eta_i = 0, \forall i \in I_{00}\} \\
\mathcal{D}_2 &= \{(d, \xi, \eta) \in \mathbb{R}^{n+2|I_{00}|} \mid \xi \geq 0, \eta \geq 0, \xi^T \eta = 0\}.
\end{aligned}$$

Como $(d^*, \xi^*, \eta^*) = (0, 0, 0)$ é um minimizador de (2.23) então novamente pelo Lema 1.7 (c)

$$(\nabla f(x^*), 0, 0) \in \mathcal{T}((0, 0, 0), \mathcal{D})^*$$

o que significa pela Proposição 2.23 que

$$(-\nabla f(x^*), 0, 0) \in \hat{N}((0, 0, 0), \mathcal{D}) \subseteq N((0, 0, 0), \mathcal{D}).$$

Por outro lado consideremos a multifunção $\Phi : \mathbb{R}^{n+2|I_{00}|} \rightrightarrows \mathbb{R}^{n+2|I_{00}|}$ dada por

$$\Phi(v) = \{w \in \mathcal{D}_1 : v + w \in \mathcal{D}_2\}$$

então como

$$\begin{aligned}
gph\Phi = \{ & (d^v, \xi^v, \eta^v, d^w, \xi^w, \eta^w) \in \mathbb{R}^{2(n+2|I_{00}|)} \mid \nabla g_i(x^*)^T d^w \leq 0, \forall i \in \mathcal{I}_g, \\
& \nabla h_i(x^*)^T d^v = 0, \forall i = 1, \dots, p, \\
& \nabla G_i(x^*)^T d^v = 0, \forall i \in I_{0+}, \\
& \nabla H_i(x^*)^T d^v = 0, \forall i \in I_{+0}, \\
& \nabla G_i(x^*)^T d^v - \xi_i^v = 0, \forall i \in I_{00}, \\
& \nabla H_i(x^*)^T d^v - \eta_i^v = 0, \forall i \in I_{00}, \\
& \xi^v + \xi^w \geq 0, \eta^v + \eta^w \geq 0, (\xi^v + \xi^w)^T (\eta^v + \eta^w) = 0 \}
\end{aligned}$$

podemos expressar $gph\Phi$ como

$$\begin{aligned}
gph\Phi = \bigcup_{(\nu_1, \nu_2) \in \mathcal{P}(1, \dots, |I_{00}|)} \{ & (d^v, \xi^v, \eta^v, d^w, \xi^w, \eta^w) \in \mathbb{R}^{2(n+2|I_{00}|)} \mid \nabla g_i(x^*)^T d^w \leq 0, \forall i \in \mathcal{I}_g \\
& \nabla h_i(x^*)^T d^v = 0, \forall i = 1, \dots, p \\
& \nabla G_i(x^*)^T d^v = 0, \forall i \in I_{0+} \\
& \nabla H_i(x^*)^T d^v = 0, \forall i \in I_{+0} \\
& \nabla G_i(x^*)^T d^v - \xi_i^v = 0, \forall i \in I_{00} \\
& \nabla H_i(x^*)^T d^v - \eta_i^v = 0, \forall i \in I_{00} \\
& \xi_{\nu_1}^v + \xi_{\nu_1}^w = 0, \xi_{\nu_2}^v + \xi_{\nu_2}^w \geq 0 \\
& \eta_{\nu_1}^v + \eta_{\nu_1}^w \geq 0, \eta_{\nu_2}^v + \eta_{\nu_2}^w = 0 \}.
\end{aligned}$$

Sendo cada um dos elementos da reunião um conjunto poliédrico convexo, temos pela Proposição 2.29 que

$$(-\nabla f(x^*), 0, 0) \in N((0, 0, 0), \mathcal{D}_1) + N((0, 0, 0), \mathcal{D}_2).$$

Como \mathcal{D}_1 é convexo então pela Proposição 2.24 seu cone normal limite coincide com o cone normal de Fréchet, logo pela Proposição 2.23

$$N((0, 0, 0), \mathcal{D}_1) = -\mathcal{D}_1^*.$$

Como

$$\begin{aligned}
\mathcal{D}_1 = \{ & (d, \xi, \eta) \in \mathbb{R}^{n+2|I_{00}|} \mid \nabla g_i(x^*)^T d \leq 0, \forall i \in \mathcal{I}_g \\
& \nabla h_i(x^*)^T d = 0, \forall i = 1, \dots, p \\
& \nabla G_i(x^*)^T d = 0, \forall i \in I_{0+} \\
& \nabla H_i(x^*)^T d = 0, \forall i \in I_{+0} \\
& \nabla G_i(x^*)^T d - \xi_i = 0, \forall i \in I_{00} \\
& \nabla H_i(x^*)^T d - \eta_i = 0, \forall i \in I_{00} \}
\end{aligned}$$

pelo Lema 1.9 existem multiplicadores λ^g , λ^h , λ^G e λ^H com $\lambda_{\mathcal{I}_g}^g \geq 0$ tal que

$$\begin{aligned}
(-\nabla f(x^*), 0, 0) \in & \sum_{i \in \mathcal{I}_g} \lambda_i^g (\nabla g_i(x^*), 0, 0) + \sum_{i=1}^p \lambda_i^h (\nabla h_i(x^*), 0, 0) \\
& - \sum_{i \in I_{0+}} \lambda_i^G (\nabla G_i(x^*), 0, 0) - \sum_{i \in I_{+0}} \lambda_i^H (\nabla H_i(x^*), 0, 0) \\
& - \sum_{i \in I_{00}} \lambda_i^G (\nabla G_i(x^*), -e_i, 0) - \sum_{i \in I_{00}} \lambda_i^H (\nabla H_i(x^*), 0, -e_i) \\
& + N((0, 0, 0), \mathcal{D}_2)
\end{aligned}$$

onde e_i é o vetor unitário na posição i em I_{00} .

Por outro lado como $\mathcal{D}_2 = \{(d, \xi, \eta) : \xi \geq 0, \eta \geq 0, \xi^T \eta = 0\}$ então pela Proposição 2.25,

$$\begin{aligned}
N((0, 0, 0), \mathcal{D}_2) &= N(0, \mathbb{R}^n) \times N((0, 0), \{(\xi, \eta) : \xi \geq 0, \eta \geq 0, \xi^T \eta = 0\}) \\
&= \{0\} \times N((0, 0), \{(\xi, \eta) : \xi \geq 0, \eta \geq 0, \xi^T \eta = 0\})
\end{aligned}$$

de onde

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i^g \nabla g_i(x^*) + \sum_{i=1}^p \lambda_i^h \nabla h_i(x^*) - \sum_{i \in I_{0+} \cup I_{00}} \lambda_i^G \nabla G_i(x^*) - \sum_{i \in I_{+0} \cup I_{00}} \lambda_i^H \nabla H_i(x^*) = 0$$

e

$$-(\lambda_{I_{00}}^G, \lambda_{I_{00}}^H) \in N((0, 0), \{(\xi, \eta) : \xi \geq 0, \eta \geq 0, \xi^T \eta = 0\}).$$

Assim pela Proposição 2.26 temos que

$$(\lambda_i^G > 0 \text{ e } \lambda_i^H > 0) \text{ ou } \lambda_i^G \lambda_i^H = 0 \quad \forall i \in I_{00}.$$

Finalmente, podemos escolher $\lambda_{I_{+0}}^G = 0$, $\lambda_{I_{+0}}^H > 0$ e $\lambda_i^g = 0$ se $i \notin \mathcal{I}_g$, para concluir que x^* é M-estacionário. \square

Devemos enfatizar que considerar MPEC-GCQ não é a única forma de conseguir

M-estacionariedade, tendo outras diferentes CQs que produzem o mesmo resultado [61, 74]. Em [61] encontramos um diagrama que mostra diferentes CQs que implicam M-estacionariedade (Figura 2.6).

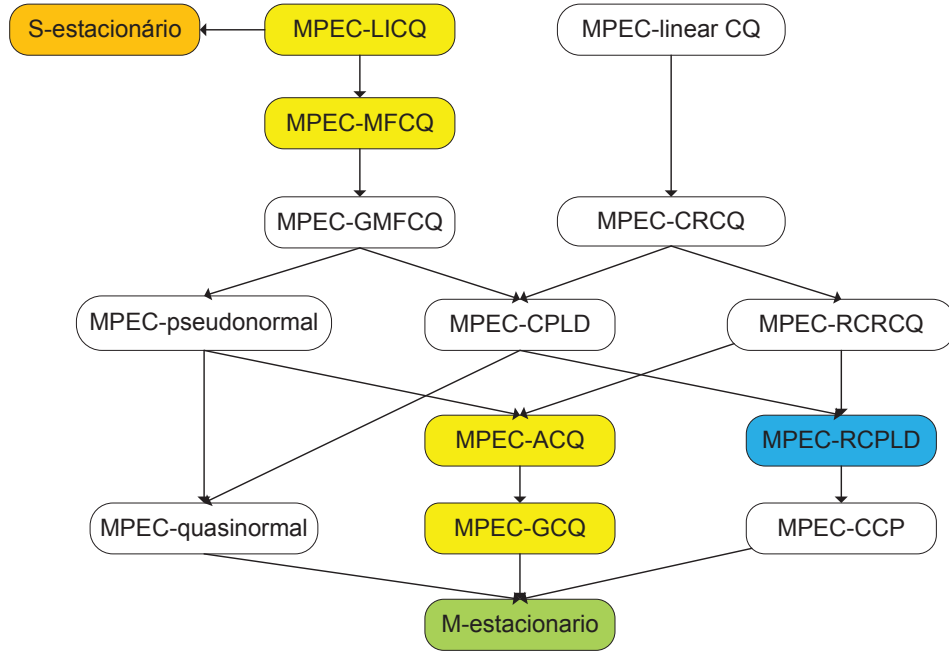


Figura 2.6: Relação entre diferentes CQs que implicam M-estacionariedade [61].

Deste grupo de CQs, além das vistas no capítulo anterior, há uma que é de nosso interesse para a análise posterior, a qual definimos a seguir.

Definição 2.31. [34, Definição 4.3] Sejam $\mathcal{I} \subset \{1, \dots, p\}$, $\mathcal{I}_G \subset I_{0+}$ e $\mathcal{I}_H \subset I_{+0}$ tais que

$$\{\nabla G_i(x^*), \nabla H_j(x^*), \nabla h_k(x^*) \mid i \in \mathcal{I}_G, j \in \mathcal{I}_H, k \in \mathcal{I}\}$$

é uma base para o espaço

$$\text{span}\{\nabla G_i(x^*), \nabla H_j(x^*), \nabla h_k(x^*) \mid i \in I_{0+}, j \in I_{+0}, k = 1, \dots, p\}.$$

Dizemos que x^* satisfaz a condição de qualificação MPEC dependência linear positiva constante relaxada (MPEC-RCPLD) se existe uma vizinhança $B(x^*)$ de x^* tal que

- $\{\nabla G_i(x), \nabla H_j(x), \nabla h_k(x) \mid i \in I_{0+}, j \in I_{+0}, k = 1, \dots, p\}$ possui o mesmo posto para todo $x \in B(x^*)$;
- para cada $\tilde{\mathcal{I}}_g \subset \mathcal{I}_g$ e $\tilde{\mathcal{I}}_G, \tilde{\mathcal{I}}_H \subset I_{00}$, se existem multiplicadores $\lambda^h, \lambda^g, \lambda^G, \lambda^H$, não todos nulos, satisfazendo $(\lambda_i^G \lambda_i^H = 0 \text{ ou } \lambda_i^G, \lambda_i^H > 0) \forall i \in I_{00}$ e

$$\sum_{k \in \mathcal{I}} \lambda_k^h \nabla h_k(x^*) + \sum_{l \in \tilde{\mathcal{I}}_g} \lambda_l^g \nabla g_l(x^*) - \sum_{i \in \mathcal{I}_G \cup \tilde{\mathcal{I}}_G} \lambda_i^G \nabla G_i(x^*) - \sum_{j \in \mathcal{I}_H \cup \tilde{\mathcal{I}}_H} \lambda_j^H \nabla H_j(x^*) = 0$$

então

$$\{\nabla G_i(x), \nabla H_j(x), \nabla h_k(x), \nabla g_l(x) \mid i \in \mathcal{I}_G \cup \tilde{\mathcal{I}}_G, j \in \mathcal{I}_H \cup \tilde{\mathcal{I}}_H\}, k \in \mathcal{I}, l \in \tilde{\mathcal{I}}_g\}$$

é linearmente dependente para cada $x \in B(x^*)$.

A condição MPEC-RCPLD será necessária no Capítulo 4 para mostrar convergência para pontos S-estacionários do algoritmo desenvolvido no capítulo seguinte.

Capítulo 3

Método PQS estabilizado

No primeiros capítulos vimos os conceitos e propriedades associados a problemas MPEC. Neste introduziremos um método especial de programação quadrática sequencial estabilizada (PQSe), para a otimização não linear, na ideia de que quando aplicado a problemas MPEC, este método converge pelo menos a pontos que são M-estacionários. Métodos PQSe têm a propriedade de gerar sequências com convergência local rápida mesmo que os gradientes das restrições ativas sejam linearmente dependentes [31]. Descreveremos um algoritmo baseado no desenvolvido por P.E. Gill, K. Kungurtsev e D.P. Robinson em [30], para o qual consideraremos problemas do tipo

$$\begin{array}{ll} \min & f(x) \\ \text{s.a} & c(x) = 0 \\ & x \geq 0 \end{array} \quad (3.1)$$

sendo $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e $c : \mathbb{R}^n \rightarrow \mathbb{R}^m$ são funções duas vezes continuamente diferenciáveis, sendo possível acondicionar problemas mais gerais do tipo (1.1) com ou sem restrições de caixa a este formato. O algoritmo considera a Lagrangiana primal dual [29] como função de mérito para garantir convergência para qualquer ponto inicial; usa uma busca linear flexível considerando uma direção que provém da solução de um subproblema quadrático estritamente convexo e, quando existe, uma direção de curvatura negativa da função de mérito primal dual. Quando certas condições são satisfeitas o subproblema quadrático é resolvido mediante um sistema de equações lineares baseado em uma estimativa do conjunto ativo ótimo; possui propriedades de convergência de primeira e segunda ordem e convergência local rápida numa vizinhança da solução.

Neste capítulo veremos a fundamentação do método bem como uma descrição do algoritmo e suas propriedades de convergência.

3.1 Conceitos preliminares

Nesta seção veremos alguns conceitos e notações próprias relacionadas com o Problema (3.1). Usaremos a notação do artigo [30] para uma justificativa mais clara dos resultados. Assim $\|\cdot\|$ denota a norma Euclidiana ou sua norma matricial induzida. O menor autovalor de uma matriz simétrica A será denotado por $\lambda_{\min}(A)$. Dado os vetores a e b , o par (a, b) denota o vetor $(a^T, b^T)^T$; se a e b possuem a mesma dimensão, o vetor com i -ésima componente $a_i b_i$ é denotado por $a \cdot b$. Similarmente, $\min(a, b)$ é um vetor com componentes $\min(a_i, b_i)$. Os vetores e e e_j representam, respectivamente, o vector coluna de uns e a j -ésima coluna da matriz identidade I . A i -ésima componente de um vetor será designada por $[\cdot]_i$, por exemplo, $[v]_i$ é a i -ésima componente do vetor v . Para um vetor dado $u \in \mathbb{R}^l$ e um conjunto de índices S , $[u]_S$ denota o subvetor de componentes u_j tal que $j \in S \cap \{1, 2, \dots, l\}$. Da mesma forma, se M é uma matriz $l \times l$ simétrica, então $[M]_S$ denota a matriz simétrica com elementos m_{ij} para $i, j \in S \cap \{1, 2, \dots, l\}$.

$J(x)$ denotará a matriz Jacobiana das restrições com linhas $\nabla c_i(x)^T$. A Lagrangiana associada a (3.1) é dada por

$$L(x, y, w) = f(x) - c(x)^T y - w^T x$$

onde y e w são vetores de variáveis duais associadas às restrições de igualdade e desigualdade, respectivamente. A Hessiana da Lagrangiana com respeito a x é denotada por $H(x, y) = \nabla^2 f(x) - \sum_{i=1}^m y_i \nabla^2 c_i(x)$.

O próximo lema caracteriza os pontos KKT o Problema (3.1) e o método busca encontrar tais pontos.

Lema 3.1. *Um ponto KKT para o Problema (3.1) é um ponto x^* para o qual existe um vetor dual y^* tal que $r(x^*, y^*) = 0$, sendo*

$$r(x, y) = \|(c(x), \min(x, \nabla f(x) - J(x)^T y))\|. \quad (3.2)$$

Demonstração. Primeiramente notemos que dados $x, u \in \mathbb{R}^n$ e definindo w como

$$w_i = \begin{cases} u_i - x_i, & \text{se } x_i \leq u_i \\ 0, & \text{se } x_i > u_i \end{cases}, \quad (3.3)$$

temos

$$\min(x, u) = u - w. \quad (3.4)$$

Se $r(x^*, y^*) = 0$, então

$$\begin{aligned} c(x^*) &= 0, \\ \min(x^*, \nabla f(x^*) - J(x^*)^T y^*) &= 0 \end{aligned}$$

e logo, considerando em (3.3) $x = x^*$, $u = \nabla f(x^*) - J(x^*)^T y^*$ e $w = w^*$, temos por (3.4)

$$\nabla f(x^*) - J(x^*)^T y^* - w^* = 0$$

e se $w_i^* \neq 0$ então por (3.3) $x_i^* = [\nabla f(x^*) - J(x^*)^T y^*]_i - w_i^* = 0$ e assim

$$x_i^* w_i^* = 0$$

além de que por construção em (3.3) $w_i^* \geq 0$.

Reciprocamente, se (x^*, y^*, z^*) satisfaz as condições KKT, então

$$\begin{aligned} \nabla f(x^*) - J(x^*)^T y^* - w^* &= 0 \\ c(x^*) &= 0 \\ x_i^* w_i^* &= 0 \\ x_i^*, w_i^* &\geq 0. \end{aligned}$$

Assim,

$$\min(x_i, [\nabla f(x) - J(x)^T y]_i) = 0$$

e portanto, $r(x^*, y^*) = 0$. □

Note que para qualquer ponto KKT (x^*, y^*, w^*) para o Problema (3.1), temos

$$w^* = \nabla f(x^*) - J(x^*)^T y^*.$$

Assim, podemos focar apenas no par (x^*, y^*) .

Qualquer (x^*, y^*) satisfazendo $r(x^*, y^*) = 0$, é chamado um *par KKT de primeira ordem*. Se houver uma condição de qualificação, então um minimizador local x^* para o Problema (3.1) é um ponto KKT de primeira ordem. Neste caso, qualquer y^* associado a um par KKT (x^*, y^*) é um vetor de multiplicadores de Lagrange para as restrições $c(x) = 0$. Dado um ponto KKT de primeira ordem x^* , o conjunto não vazio de vetores duais $\mathcal{Y}(x^*)$ se define como

$$\mathcal{Y}(x^*) = \{y \in \mathbb{R}^m : (x^*, y) \text{ satisfaz } r(x^*, y) = 0\}. \quad (3.5)$$

Dado $x \geq 0$, o conjunto ativo em x será denotado por

$$\mathcal{A}(x) = \{i : [x]_i = 0\}. \quad (3.6)$$

3.2 Fundamentação do método

A ideia é juntar duas formas de abordagem do Problema (3.1). A primeira é o método de Programação Quadrática Sequencial estabilizado, que envolve a solução de um subproblema de programação quadrática (PQ), no qual um modelo quadrático local da Lagrangiana é minimizado sujeito às restrições linearizadas [29]. Dada uma estimativa (x_k, y_k) da solução primal-dual (x^*, y^*) do Problema (3.1) o método PQS estabilizado encontra uma nova estimativa resolvendo o subproblema

$$\begin{aligned} \min \quad & \nabla f(x_k)^T(x - x_k) + \frac{1}{2}(x - x_k)^T H_k(x - x_k) + \frac{1}{2}\mu_k^R \|y\|^2 \\ \text{s.a} \quad & c(x_k) + J(x_k)(x - x_k) + \mu_k^R(y - y_k) = 0 \\ & x \geq 0 \end{aligned} \quad (3.7)$$

onde $H_k = H(x_k, y_k)$ e $\mu_k^R \rightarrow 0$ ($\mu_k^R > 0$). O fato de que $\mu_k^R \rightarrow 0$, está associado a que quando $\mu_k^R = 0$ então o subproblema (3.7) é o subproblema quadrático do método PQS usual.

A segunda forma baseia-se na função Lagrangiana aumentada primal-dual [29]

$$M(x, y; y^E, \mu) = f(x) - c(x)^T y^E + \frac{1}{2\mu} \|c(x)\|^2 + \frac{1}{2\mu} \|c(x) + \mu(y - y^E)\|^2 \quad (3.8)$$

onde μ é um parâmetro de penalidade positivo e y^E é uma estimativa de multiplicador de Lagrange. O seguinte teorema motiva o uso de (3.8) como uma função de mérito e tem como hipótese a condição SOSC da Definição 1.14.

Teorema 3.2. [30, Teorema 1.3] *Se (x^*, y^*) é uma solução do Problema (3.1) que satisfaz SOSC, então para $y^E = y^*$, existe $\bar{\mu}$ positivo tal que para todo $0 < \mu < \bar{\mu}$, o ponto (x^*, y^*) cumpre as condições suficientes de otimalidade de segunda ordem para o problema*

$$\begin{aligned} \min_{x, y} \quad & M(x, y, y^E, \mu) \\ \text{s.a} \quad & x \geq 0. \end{aligned}$$

O resultado do Teorema 3.2 implica que podemos resolver o Problema (3.1), resolvendo o problema anterior. A ideia é desenhar um algoritmo em termos de uma sequência de estimativas de multiplicadores $\{y_k^E\}$, onde se possível $y_k^E = y_k$ e que considere a M como função de mérito. Via PQS se definem iterações que correspondem a minimizar um modelo quadrático da função de mérito sujeito a restrições de caixa. Cada modelo quadrático é definido em termos do gradiente de M e uma aproximação B da Hessiana de M , assim se $v_k = (x_k, y_k)$ denota a k -ésima estimativa de uma solução primal-dual do Problema (3.1), o modelo quadrático local está dado por

$$Q_k(v, y_k^E, \mu) = (v - v_k)^T \nabla M(v_k, y_k^E, \mu) + \frac{1}{2}(v - v_k)^T B(v_k, \mu)(v - v_k). \quad (3.9)$$

Para valores dados y^E e μ , o gradiente e a Hessiana de M em relação a (x, y) têm a forma

$$\begin{aligned}\nabla M(x, y, y^E, \mu) &= \begin{pmatrix} \nabla f(x) - J(x)^T y^E + \frac{1}{\mu} J(x)^T c(x) + \frac{1}{\mu} J(x)^T (c(x) + \mu(y - y^E)) \\ c(x) + \mu(y - y^E) \end{pmatrix} \\ &= \begin{pmatrix} \nabla f(x) - J(x)^T \left(\left(y^E - \frac{1}{\mu} c(x) \right) + (y^E - \frac{1}{\mu} c(x) - y) \right) \\ \mu \left(y - \left(y^E - \frac{1}{\mu} c(x) \right) \right) \end{pmatrix} \\ &= \begin{pmatrix} \nabla f(x) - J(x)^T (\pi(x, y^E, \mu) + (\pi(x, y^E, \mu) - y)) \\ \mu(y - \pi(x, y^E, \mu)) \end{pmatrix} \quad (3.10)\end{aligned}$$

onde $\pi(x, y^E, \mu) = y^E - c(x)/\mu$, e

$$\begin{aligned}\nabla^2 M(x, y, y^E, \mu) &= \begin{pmatrix} \nabla^2 f(x) - \sum_{i=1}^m (2\pi(x, y^E, \mu) - y) \nabla^2 c_i(x) + \frac{2}{\mu} J(x)^T J(x) & J(x)^T \\ J(x) & \mu I \end{pmatrix} \\ &= \begin{pmatrix} H(x, \pi(x, y^E, \mu) + (\pi(x, y^E, \mu) - y)) + \frac{2}{\mu} J(x)^T J(x) & J(x)^T \\ J(x) & \mu I \end{pmatrix} \quad (3.11)\end{aligned}$$

onde $H(x, y) = \nabla^2 f(x) - \sum_{i=1}^m y_i \nabla^2 c_i(x)$. A aproximação da Hessiana que usaremos em (3.9), resulta de substituir $\pi(x_k, y_k^E, \mu)$ por y_k na Hessiana $\nabla^2 M(x_k, y_k, y_k^E, \mu)$ [30]. Assim a matriz B em (3.9) é definida por

$$B(v_k; \mu) = \begin{pmatrix} H(x_k, y_k) + \frac{2}{\mu} J(x_k)^T J(x_k) & J(x_k)^T \\ J(x_k) & \mu I \end{pmatrix} \quad (3.12)$$

tendo a particularidade de que é independente da estimativa dos multiplicadores y_k^E . Além disso, o benefício de usar $B(v_k; \mu)$ no modelo quadrático (3.9) é que o subproblema quadrático

$$\begin{aligned} \min_v \quad & Q_k(v; y_k^E, \mu) \\ \text{s.a.} \quad & [v]_i \geq 0 \quad i = 1, \dots, n \end{aligned} \quad (3.13)$$

é equivalente, sob certas hipóteses, ao subproblema PQS estabilizado

$$\begin{aligned} \min_{x, y} \quad & \nabla f(x_k)^T (x - x_k) + \frac{1}{2} (x - x_k)^T H(x_k, y_k) (x - x_k) + \frac{1}{2} \mu \|y\|^2 \\ \text{s.a.} \quad & c(x_k) + J(x_k)(x - x_k) + \mu(y - y_k^E) = 0, \quad x \geq 0, \end{aligned}$$

como pode ser visto no seguinte resultado.

Teorema 3.3. [29, Resultado 2.1] *Seja $\mu_k^R > 0$, $v_k = (x_k, y_k)$, $c_k = c(x_k)$ e $J_k = J(x_k)$. Dada uma matriz $\hat{H}_k = \hat{H}(x_k, y_k)$ tal que $\hat{H}_k + (1/\mu_k^R) J_k^T J_k$ é definida positiva, considere*

o problema

$$\begin{aligned} \min_{x,y} \quad & \nabla f(x_k)^T(x - x_k) + \frac{1}{2}(x - x_k)^T \hat{H}_k(x - x_k) + \frac{1}{2}\mu_k^R \|y\|^2 \\ \text{s.a.} \quad & c_k + J_k(x - x_k) + \mu_k^R(y - y_k^E) = 0 \\ & x \geq 0 \end{aligned} \quad (3.14)$$

que é o subproblema PQS estabilizado (3.7) com $H_k = \hat{H}_k$ e $y_k = y_k^E$. Então:

1. O PQ estabilizado (3.14) tem uma única solução primal-dual $\hat{v}_k = (\hat{x}_k, \hat{y}_k)$;
2. A única solução $\hat{v}_k = (\hat{x}_k, \hat{y}_k)$ do PQ estabilizado (3.14) é a única solução do problema

$$\begin{aligned} \min_v \quad & \nabla M(v_k; y_k^E, \mu_k^R)^T(v - v_k) + \frac{1}{2}(v - v_k)^T \hat{B}(v_k; \mu_k^R)(v - v_k) \\ \text{s.a.} \quad & [v]_i \geq 0 \quad i = 1, 2, \dots, n \end{aligned} \quad (3.15)$$

onde

$$\hat{B}(v_k; \mu_k^R) = \begin{pmatrix} \hat{H}(x_k, y_k) + \frac{2}{\mu_k^R} J(x_k)^T J(x_k) & J(x_k)^T \\ J(x_k) & \mu_k^R I \end{pmatrix}.$$

Esta equivalência sugere um algoritmo baseado na solução de (3.13) com Q_k definido com um pequeno valor $\mu = \mu_k^R$ pois $\mu_k^R \rightarrow 0$. Porém se $M(x, y; y_k^E, \mu)$ (3.8) serve como uma função de mérito para fazer a busca linear, ela deve ser definida com uma penalidade μ não muito pequena. Assim, são considerados diferentes parâmetros μ_k^R e μ_k , com μ_k^R utilizado para a definição do modelo quadrático local e μ_k utilizado para a busca linear da função de mérito.

Note que no teorema anterior a matriz \hat{B} precisa ser definida positiva, fato que é garantido pelo seguinte lema no qual se usa a inércia de uma matriz simétrica A , isto é,

$$In(A) = (n_+, n_-, n_0)$$

sendo n_+ , n_- e n_0 são respectivamente o número de autovalores positivos, negativos e nulos da matriz A .

Lema 3.4. [29, Lema 2.2] *Sejam $\mu > 0$, H simétrica de ordem $n \times n$ e J de ordem $m \times n$. Se definimos*

$$B = \begin{pmatrix} H + \frac{2}{\mu} J^T J & J^T \\ J & \mu I \end{pmatrix} \quad e \quad K = \begin{pmatrix} H & J^T \\ J & -\mu I \end{pmatrix}$$

então

1. A matriz $H + \frac{1}{\mu} J^T J$ é definida positiva se e somente se $In(B) = (n + m, 0, 0)$;

2. A matriz $H + \frac{1}{\mu} J^T J$ é definida positiva se e somente se $\text{In}(K) = (n, m, 0)$.

Assim, para verificar se B é definida positiva, somente basta analisar a inércia da matriz K . O termo não-zero μI em B serve para regularizar o sistema de equações associado a B , isto é, não é necessário que J tenha posto completo para garantir que o sistema seja não singular. Além disso, se $\mu = 0$ então K é a matriz KKT do método SQP usual. Isso motiva a considerar $\mu = \mu_k^R$ pequeno como um parâmetro de regularização.

3.3 Descrição do Algoritmo

O Algoritmo segue os seguintes passos:

1. Encontra uma direção de curvatura negativa $s_k^{(1)}$ de Q_k (expressão (3.9)) e uma estimativa ξ_k do menor autovalor de $B_{\mathcal{F}_\epsilon}$;
2. Atualiza o multiplicador estimado y_k^E e o parâmetro de regularização μ_k^R ;
3. Encontra a direção para a busca linear;
4. Realiza uma busca linear flexível e atualiza o parâmetro de penalidade μ_k .

Em cada iteração são conhecidos (x_k, y_k) , μ_{k-1} e μ_{k-1}^R . Para atualizar y_k^E e μ_k^R para a nova iteração, é necessário considerar o conjunto $\mathcal{A}(x)$ definido em (3.6), definindo-se o conjunto primal-dual livre em x

$$\mathcal{F}(x) = \{1, 2, \dots, n + m\} \setminus \mathcal{A}(x) \quad (3.16)$$

e os conjuntos ϵ -ativo e ϵ -livre definidos respectivamente por

$$\mathcal{A}_\epsilon(x, y, \mu) = \{i : x_i \leq \epsilon\} \text{ e} \quad (3.17)$$

$$\mathcal{F}_\epsilon(x, y, \mu) = \{1, 2, \dots, n + m\} \setminus \mathcal{A}_\epsilon(x, y, \mu) \quad (3.18)$$

onde $\epsilon = \min(\epsilon_a, \max(\mu, r(x, y)^{\bar{\gamma}}))$ para $0 < \epsilon_a, \bar{\gamma} < 1$ fixos.

Quando μ é suficientemente pequeno e (x, y) está próximo a um par KKT (x^*, y^*) então $\mathcal{F}(x^*) \subset \mathcal{F}_\epsilon(x, y, \mu)$. De fato se $i \notin \mathcal{A}(x^*)$ então $x_i^* > 0$, além disso, pelo Lema 3.1 $r(x^*, y^*) = 0$, o que implica que $r(x, y)$ esta próximo de zero. Logo podemos considerar μ satisfazendo

$$\begin{aligned} r(x, y)^{\bar{\gamma}} &\leq \mu < x_i^*, \\ |x_i - x_i^*| &< x_i^* - \mu \text{ e} \\ \mu &\leq \epsilon_a, \end{aligned}$$

assim $\mathcal{A}_\epsilon(x, y, \mu) = \{i : x_i \leq \mu\}$ e $x_i > \mu$, de onde $i \notin \mathcal{A}_\epsilon(x, y, \mu)$, o que implica $\mathcal{F}(x^*) \subset \mathcal{F}_\epsilon(x, y, \mu)$.

A análise anterior implica que para μ suficientemente pequeno $\mathcal{A}_\epsilon(x, y, \mu) \subset \mathcal{A}(x^*)$ o que é de esperar quando estamos próximos a um ponto KKT. Isto permite desenhar o algoritmo considerando somente o conjunto $\mathcal{F}_\epsilon(x, y, \mu)$, já que se $i \in \mathcal{A}_\epsilon(x, y, \mu)$, então $x_i = 0$.

3.3.1 Estimando uma direção de curvatura negativa de Q_k

Calculado $\mathcal{F}_\epsilon = \mathcal{F}_\epsilon(x_k, y_k, \mu_{k-1}^R)$, y_k^E e μ_k^R dependem da estimativa do menor autovalor de $B_{\mathcal{F}_\epsilon}(v_k; \mu_{k-1}^R)$ computado das ϵ -livre colunas da Hessiana aproximada $B(v_k; \mu_{k-1}^R)$ dada em (3.12). Pelo Lema 3.4 (1), esta aproximação é feita em termos de uma estimativa do menor autovalor de $H_{\mathcal{F}_\epsilon} + \frac{1}{\mu_{k-1}^R} J_{\mathcal{F}_\epsilon}^T J_{\mathcal{F}_\epsilon}$ o qual denotaremos por

$$\lambda_{\min} = \lambda_{\min}(H_{\mathcal{F}_\epsilon} + \frac{1}{\mu_{k-1}^R} J_{\mathcal{F}_\epsilon}^T J_{\mathcal{F}_\epsilon}),$$

onde $H_{\mathcal{F}_\epsilon}$ é a matriz de ϵ -livre linhas e colunas de $H(x_k, y_k)$ e $J_{\mathcal{F}_\epsilon}$ é a matriz de ϵ -livre colunas de $J(x_k)$. Independentemente do algoritmo, devemos ter uma maneira de determinar se $H_{\mathcal{F}_\epsilon} + (1/\mu_{k-1}^R) J_{\mathcal{F}_\epsilon}^T J_{\mathcal{F}_\epsilon}$ é semidefinida positiva, caso contrário, ter como encontrar uma direção $u_{\mathcal{F}_\epsilon}$ satisfazendo

$$u_{\mathcal{F}_\epsilon}^T (H_{\mathcal{F}_\epsilon} + \frac{1}{\mu_{k-1}^R} J_{\mathcal{F}_\epsilon}^T J_{\mathcal{F}_\epsilon}) u_{\mathcal{F}_\epsilon} \leq \theta \lambda_{\min}(H_{\mathcal{F}_\epsilon} + \frac{1}{\mu_{k-1}^R} J_{\mathcal{F}_\epsilon}^T J_{\mathcal{F}_\epsilon}) \|u_{\mathcal{F}_\epsilon}\|^2 < 0 \quad (3.19)$$

onde $\theta > 0$ é independente de x_k e y_k . O algoritmo encontra uma direção aproximada de curvatura negativa $s_k^{(1)}$ da função de mérito e uma estimativa ξ_k não negativa de $\max(0, -\lambda_{\min})$.

Se $H_{\mathcal{F}_\epsilon} + \frac{1}{\mu_{k-1}^R} J_{\mathcal{F}_\epsilon}^T J_{\mathcal{F}_\epsilon} > 0$ então $\xi_k = 0$ e $s_k^{(1)} = 0$.

$$\text{Caso contrário} \begin{cases} \xi_k = -u_{\mathcal{F}_\epsilon}^T (H_{\mathcal{F}_\epsilon} + (1/\mu_{k-1}^R) J_{\mathcal{F}_\epsilon}^T J_{\mathcal{F}_\epsilon}) u_{\mathcal{F}_\epsilon} / \|u_{\mathcal{F}_\epsilon}\|^2 \\ \text{encontra } s_k^{(1)} \text{ tal que } s_k^{(1)T} B(v_k; \mu_{k-1}^R) s_k^{(1)} < 0 \end{cases}$$

Algoritmo 1 (estima uma direção de curvatura negativa $s_k^{(1)}$)

- 1: **Dados** $x_k, y_k, \mu_{k-1}^R, J_k, H_k$
- 2: Computar $H_{\mathcal{F}_\epsilon}$ e $J_{\mathcal{F}_\epsilon}$ como submatrizes de H_k e J_k associadas a $\mathcal{F}_\epsilon(x_k, y_k, \mu_{k-1}^R)$;
- 3: **se** $H_{\mathcal{F}_\epsilon} + \frac{1}{\mu_{k-1}^R} J_{\mathcal{F}_\epsilon}^T J_{\mathcal{F}_\epsilon}$ é positiva semidefinida **então**
- 4: $\xi_k = 0, u_k^{(1)} = 0$ e $w_k^{(1)} = 0$;
- 5: **caso contrário**
- 6: Computar $u_{\mathcal{F}_\epsilon} \neq 0$ que cumpra (3.19);

```

7:           $\xi_k = -u_{\mathcal{F}_\epsilon}^T (H_{\mathcal{F}_\epsilon} + (1/\mu_{k-1}^R) J_{\mathcal{F}_\epsilon}^T J_{\mathcal{F}_\epsilon}) u_{\mathcal{F}_\epsilon} / \|u_{\mathcal{F}_\epsilon}\|^2 > 0;$ 
8:           $u_k^{(1)} = 0;$   $[u_k^{(1)}]_{\mathcal{F}_\epsilon} = u_{\mathcal{F}_\epsilon};$  e
9:           $w_k^{(1)} = -(1/\mu_{k-1}^R) J_k u_k^{(1)};$ 
10:      fim
11:       $s_k^{(1)} = (u_k^{(1)}, w_k^{(1)});$ 
12:      saída  $(s_k^{(1)}, \xi_k);$ 
13: fim

```

O seguinte Lema mostra as propriedades de ξ_k e $s_k^{(1)}$ obtidas no Algoritmo 1:

Lema 3.5. [30, Lema 2.1] *Suponha que ξ_k e $s_k^{(1)} = (u_k^{(1)}, w_k^{(1)})$ são obtidas no Algoritmo 1, e que a matriz $H_{\mathcal{F}_\epsilon} + \frac{1}{\mu_{k-1}^R} J_{\mathcal{F}_\epsilon}^T J_{\mathcal{F}_\epsilon}$ não é semidefinida positiva. Então*

1. $u_k^{(1)} \neq 0$, $s_k^{(1)} \neq 0$, $\xi_k > 0$ e
2. $s_k^{(1)}$ é uma direção de curvatura negativa para $Q_k(v; y_{k-1}^E, \mu_{k-1}^R)$, o modelo quadrático (3.9). Em particular $s_k^{(1)}$ satisfaz

$$s_k^{(1)T} B(v_k; \mu_{k-1}^R) s_k^{(1)} \leq \theta \lambda_{\min}(H_{\mathcal{F}_\epsilon} + \frac{1}{\mu_{k-1}^R} J_{\mathcal{F}_\epsilon}^T J_{\mathcal{F}_\epsilon}) \left\| u_k^{(1)} \right\|^2 < 0,$$

onde θ é o valor considerado em (3.19).

3.3.2 Atualizando o multiplicador estimado y_k^E e o parâmetro de regularização μ_k^R

Primeiramente lembremos que em algoritmos baseados em Lagrangiano Aumentado [3, Algoritmo 3.1], a atualização do multiplicador de uma restrição de desigualdade g_i tem a forma

$$[y_k]_i = \max \{0, [\bar{y}_{k-1}]_i + \rho_{k-1} g_i(x_k)\} \quad (3.20)$$

onde ρ_{k-1} é o parâmetro de penalização e \bar{y}_{k-1} é a projeção de y_{k-1} sobre uma caixa $[-y_{\max}, y_{\max}]$.

Pelo visto no Teorema 3.2 a ideia é definir se possível $y_k^E = y_k$. Como estamos interessados em analisar as propriedades de convergência do algoritmo quando aplicado a problemas MPEC, encontraremos na Seção 4.2 que a atualização de $[y_k]_i$ neste caso tem alguma semelhança com a atualização (3.20). Na Seção 4.2 também descobriremos que y_k^E desempenha um papel semelhante ao \bar{y}_{k-1} em (3.20), por causa dessa semelhança e diferentemente do algoritmo proposto em [30], pediremos que y_k^E seja limitado.

O multiplicador estimado y_k^E será atualizado como y_k quando (x_k, y_k) melhora pelo menos uma de duas funções de mérito que medem a precisão de (x_k, y_k) como uma

estimativa de (x^*, y^*) ; estas funções $\eta(x_k)$ e $\omega(x_k, y_k, \xi_k)$ são medidas de violação da viabilidade e da otimalidade

$$\eta(x_k) = \|c(x_k)\| \quad \text{e} \quad \omega(x_k, y_k, \xi_k) = \max(\|\min(x_k, \nabla f(x_k) - J(x_k)^T y_k)\|, \xi_k). \quad (3.21)$$

Para $0 < \beta \ll 1$ considere as funções de mérito

$$\phi_V(x_k, y_k, \xi_k) = \eta(x_k) + \beta\omega(x_k, y_k, \xi_k) \quad \text{e} \quad \phi_O(x_k, y_k, \xi_k) = \beta\eta(x_k) + \omega(x_k, y_k, \xi_k). \quad (3.22)$$

Estas duas funções medem a otimalidade e a viabilidade simultaneamente. ϕ_V dá mais peso à viabilidade e ϕ_O , à otimalidade. Para a análise de convergência, esse peso não é importante. Porém, como na prática, o que mais importa é a viabilidade, ϕ_V terá um papel mais relevante na análise da convergência.

Com relação à atualização de y_k^E :

- Consideramos $y_k^E = y_k$ se $v_k = (x_k, y_k)$ satisfaz

$$\|y_k\| \leq y_{\max} \quad \text{e} \quad (3.23)$$

$$\phi_V(v_k, \xi_k) \leq \frac{1}{2}\phi_{V,k-1}^{\max} \quad \text{ou} \quad \phi_O(v_k, \xi_k) \leq \frac{1}{2}\phi_{O,k-1}^{\max}. \quad (3.24)$$

onde ϕ_V^{\max} e ϕ_O^{\max} são limitantes atualizados ao longo do processo de solução. Neste caso, a iteração k é chamada de *V-iteração* ou *O-iteração*. Além disso são atualizados:

$$\phi_{V,k}^{\max} = \frac{1}{2}\phi_{V,k-1}^{\max} \quad \text{ou} \quad \phi_{O,k}^{\max} = \frac{1}{2}\phi_{O,k-1}^{\max} \quad (3.25)$$

$$\tau_k = \frac{1}{2}\tau_{k-1}$$

$$\mu_k^R = \begin{cases} \min(\mu_0^R, \max(r_k, \xi_k)^{\bar{\gamma}}) & \text{se } \max(r_k, \xi_k) > 0 \\ \frac{1}{2}\mu_{k-1}^R & \text{no outro caso} \end{cases} \quad (3.26)$$

com $r_k = r(x_k, y_k)$ e $\bar{\gamma} \in (0, 1)$ fixo.

Este conjunto de iterações será fundamental na análise de convergência, já que se temos uma subsequência convergente, então $\phi_V(x_k, y_k, \xi_k) \rightarrow 0$ ou $\phi_O(x_k, y_k, \xi_k) \rightarrow 0$ o que implica que $\eta(x_k) \rightarrow 0$ e $\omega(x_k, y_k, \xi_k) \rightarrow 0$. Logo pelo Lema 3.1 o limite (x^*, y^*) será um par KKT, já que neste caso $r(x^*, y^*) = 0$. Também enfatizamos que neste caso (3.23) implica que y_k^E é limitado, o que será fundamental ao analisar a convergência do algoritmo em problemas MPEC.

- Se as condições de ser *V-O-iteração* não são satisfeitas, verificamos se (x_k, y_k) é uma

solução aproximada de segunda ordem do problema

$$\begin{aligned} \min_{x,y} \quad & M(x, y, y_{k-1}^E, \mu_{k-1}^R) \\ \text{s.a.} \quad & x \geq 0 \end{aligned} \quad (3.27)$$

isto é, verificamos que

$$\begin{aligned} \|\min(x_k, \nabla_x M(x_k, y_k; y_{k-1}^E, \mu_{k-1}^R))\| &\leq \tau_{k-1} \\ \|\nabla_y M(x_k, y_k; y_{k-1}^E, \mu_{k-1}^R)\| &\leq \tau_{k-1} \mu_{k-1}^R \\ \xi_k &\leq \tau_{k-1}, \end{aligned} \quad (3.28)$$

para alguma pequena tolerância τ_{k-1} . Se as condições (3.28) forem satisfeitas, a iteração se chama *M-iteração* e y_{k-1}^E é atualizada com a estimativa projetada como no algoritmo de Lagrangiano Aumentado [3],

$$y_k^E = \max(-y_{\max}e, \min(y_k, y_{\max}e)), \quad (3.29)$$

onde e é o vetor formado por uns, além disso são atualizados

$$\begin{aligned} \tau_k &= \frac{1}{2} \tau_{k-1} \\ \mu_k^R &= \begin{cases} \min(\frac{1}{2} \mu_{k-1}^R, \max(r_k, \xi_k)^{\tilde{\gamma}}) & \text{se } \max(r_k, \xi_k) > 0 \\ \frac{1}{2} \mu_{k-1}^R & \text{outro caso.} \end{cases} \end{aligned} \quad (3.30)$$

- Finalmente, no caso de que as condições de ser *V-O-M-iteração* não valer então $y_k^E = y_{k-1}^E$, $\mu_k^R = \mu_{k-1}^R$, $\phi_{V,k}^{\max} = \phi_{V,k-1}^{\max}$, $\phi_{O,k}^{\max} = \phi_{O,k-1}^{\max}$ e $\tau_k = \tau_{k-1}$ e a iteração é chamada de *F-iteração*.

3.3.3 Definição da direção de busca linear

Uma vez computados y_k^E e μ_k^R , duas direções de busca linear são encontradas: d_k relacionada com a solução de um problema quadrático e s_k associada à direção de curvatura negativa.

Resolver o Problema (3.13) com $\mu = \mu_k^R$, isto é,

$$\begin{aligned} \min_v \quad & Q_k(v; y_k^E, \mu_k^R) = (v - v_k)^T \nabla M(v_k; y_k^E, \mu_k^R) + \frac{1}{2} (v - v_k)^T B(v_k; \mu_k^R) (v - v_k) \\ \text{s.a} \quad & [v]_i \geq 0 \quad i = 1, \dots, n \end{aligned} \quad (3.31)$$

não é possível em geral, pois a matriz $B(v_k; \mu_k^R)$ nem sempre é definida positiva. Esta dificuldade é superada considerando uma aproximação do subproblema (3.31) pelo problema

convexo

$$\begin{aligned} \min_v \quad & \hat{Q}_k(v; y_k^E, \mu_k^R) = (v - v_k)^T \nabla M(v_k; y_k^E, \mu_k^R) + \frac{1}{2}(v - v_k)^T \hat{B}(v_k; \mu_k^R)(v - v_k) \\ \text{s.a} \quad & [v]_i \geq 0 \quad i = 1, \dots, n \end{aligned} \quad (3.32)$$

onde $\hat{B}(v_k; \mu_k^R)$ é uma aproximação definida positiva de $B(v_k; \mu_k^R)$ que tem a forma

$$\hat{B}(v_k; \mu_k^R) = \begin{pmatrix} \hat{H}(x_k, y_k) + 2J(x_k)^T J(x_k) & J(x_k)^T \\ J(x_k) & \mu_k^R I \end{pmatrix} \quad (3.33)$$

e $\hat{H}(x_k, y_k)$ é definida de modo que $\hat{H}(x_k, y_k) + (1/\mu_k^R)J(x_k)^T J(x_k)$ seja definida positiva (o que garante a positividade de $\hat{B}(v_k; \mu_k^R)$ pelo Lema 3.4). Assim $\hat{B}_{\mathcal{F}_\epsilon}(v_k; \mu) = B_{\mathcal{F}_\epsilon}(v_k; \mu)$ é definida positiva se $B_{\mathcal{F}_\epsilon}(v_k; \mu)$ é definida positiva.

A análise anterior justifica considerar dois casos para a direção d_k :

- No caso em que $B_{\mathcal{F}_\epsilon} > 0$ e se tenha uma V - O iteração ($y_k^E = y_k$) então definimos uma direção $d_k = \hat{v}_k - v_k$ chamada *direção de descida local* onde \hat{v}_k é a única solução do subproblema

$$\begin{aligned} \min_v \quad & Q_k(v; y_k^E, \mu_k^R) = (v - v_k)^T \nabla M(v_k; y_k^E, \mu_k^R) + \frac{1}{2}(v - v_k)^T B(v_k; \mu_k^R)(v - v_k) \\ \text{s.a} \quad & [v]_{\mathcal{A}_\epsilon} = 0. \end{aligned} \quad (3.34)$$

Para isso consideramos

$$\begin{aligned} [\hat{v}_k^{(0)}]_{\mathcal{A}_\epsilon} &= 0 \\ [\hat{v}_k^{(0)}]_{\mathcal{F}_\epsilon} &= [v_k]_{\mathcal{F}_\epsilon} \\ \hat{v}_k &= \hat{v}_k^{(0)} + \Delta \hat{v}_k^{(0)} \\ [\Delta \hat{v}_k^{(0)}]_{\mathcal{A}_\epsilon} &= 0 \end{aligned}$$

e $[\Delta \hat{v}_k^{(0)}]_{\mathcal{F}_\epsilon}$ solução de

$$B_{\mathcal{F}_\epsilon}[\Delta \hat{v}_k^{(0)}]_{\mathcal{F}_\epsilon} = - \left[\nabla Q_k(\hat{v}_k^{(0)}; y_k^E, \mu_k^R) \right]_{\mathcal{F}_\epsilon},$$

onde esta equação provém de aplicar KKT em (3.34):

$$\begin{aligned} & \nabla M + B(\hat{v}_k - v_k) + \sum_{i \in \mathcal{A}_\epsilon} s_i e_i = 0 \\ \Rightarrow \quad & \nabla M + B(\hat{v}_k^{(0)} - v_k) + B(\hat{v}_k - \hat{v}_k^{(0)}) + \sum_{i \in \mathcal{A}_\epsilon} s_i e_i = 0 \\ \Rightarrow \quad & \nabla Q_k(\hat{v}_k^{(0)}; y_k^E, \mu_k^R) + \sum_{i \in \mathcal{A}_\epsilon} s_i e_i = -B(\hat{v}_k - \hat{v}_k^{(0)}) \\ \Rightarrow \quad & \left[\nabla Q_k(\hat{v}_k^{(0)}; y_k^E, \mu_k^R) \right]_{\mathcal{F}_\epsilon} = -B_{\mathcal{F}_\epsilon}[\Delta \hat{v}_k^{(0)}]_{\mathcal{F}_\epsilon}. \end{aligned}$$

d_k é considerada na busca linear se

$$[v_k + d_k]_i \geq 0, i = 1, \dots, n; \quad [\nabla Q_k(v_k + d_k; y_k^E, \mu_k^R)]_{\mathcal{A}_\epsilon} \geq -t_k e \quad \text{e} \quad \nabla M_k^T d_k < 0 \quad (3.35)$$

com $t_k = (r_k)^\lambda$ e $\lambda = \min\{\bar{\gamma}, 1 - \bar{\gamma}\}$.

- Se $B_{\mathcal{F}_\epsilon}$ não for definida positiva então se define $d_k = \hat{v}_k - v_k$ chamada *direção de descida global*, onde \hat{v}_k é a única solução do subproblema (3.32). Para isso consideramos

$$\begin{aligned} [\hat{v}_k^{(0)}]_{\mathcal{A}_\epsilon} &= 0 \\ [\hat{v}_k^{(0)}]_{\mathcal{F}_\epsilon} &= [v_k]_{\mathcal{F}_\epsilon} \\ [\Delta \hat{v}_k^{(0)}]_{\mathcal{A}_\epsilon} &= 0 \end{aligned}$$

e $[\Delta \hat{v}_k^{(0)}]_{\mathcal{F}_\epsilon}$ solução de

$$\hat{B}_{\mathcal{F}_\epsilon} [\Delta \hat{v}_k^{(0)}]_{\mathcal{F}_\epsilon} = - \left[\nabla \hat{Q}_k(\hat{v}_k^{(0)}; y_k^E, \mu_k^R) \right]_{\mathcal{F}_\epsilon}.$$

Encontramos $\hat{\alpha}_0 \geq 0$ e $\hat{v}_k^{(1)}$ tal que $\hat{v}_k^{(1)} = \hat{v}_k^{(0)} + \hat{\alpha}_0 \Delta \hat{v}_k^{(0)}$ seja viável, e se resolve o PQ para \hat{v}_k considerando a $\hat{v}_k^{(1)}$ como ponto inicial.

Algoritmo 2 (encontra uma direção de descida d_k)

- 1: **Dados** $x_k, y_k, \mu_k^R, y_k^E, J_k, H_k$
 - 2: **Constantes:** $0 < \lambda < \min\{\bar{\gamma}, 1 - \bar{\gamma}\} < 1$;
 - 3: $B = B(x_k, y_k; \mu_k^R)$; Computar uma matriz definida positiva \hat{B} (3.33) a partir de B ;
 - 4: $\nabla M_k = \nabla M(x_k, y_k; y_k^E, \mu_k^R)$; $t_k = r(x_k, y_k)^\lambda$; $v_k = (x_k, y_k)$;
 - 5: $[\hat{v}_k^{(0)}]_{\mathcal{A}_\epsilon} = 0$; $[\hat{v}_k^{(0)}]_{\mathcal{F}_\epsilon} = [v_k]_{\mathcal{F}_\epsilon}$;
 - 6: **se** ($B_{\mathcal{F}_\epsilon}$ é definida positiva e v_k é uma V-O iteração) **então**
 - 7: $[\Delta \hat{v}_k^{(0)}]_{\mathcal{A}_\epsilon} = 0$; $B_{\mathcal{F}_\epsilon} [\Delta \hat{v}_k^{(0)}]_{\mathcal{F}_\epsilon} = - \left[\nabla Q_k(\hat{v}_k^{(0)}) \right]_{\mathcal{F}_\epsilon}$; $\hat{v}_k = \hat{v}_k^{(0)} + \Delta \hat{v}_k^{(0)}$;
 - 8: $d_k = \hat{v}_k - v_k$;
 - 9: **se** ($v_k + d_k$ é viável e $\nabla M_k^T d_k < 0$ e $[\nabla Q_k(v_k + d_k)]_{\mathcal{A}_\epsilon} \geq -t_k e$) **então**
 - 10: **saída** d_k ; (direção de descida local)
 - 11: **fim**
 - 12: **fim**
 - 13: $[\Delta \hat{v}_k^{(0)}]_{\mathcal{A}_\epsilon} = 0$; Resolver $\hat{B}_{\mathcal{F}_\epsilon} [\Delta \hat{v}_k^{(0)}]_{\mathcal{F}_\epsilon} = - \left[\nabla \hat{Q}_k(\hat{v}_k^{(0)}) \right]_{\mathcal{F}_\epsilon}$;
 - 14: Computar $\hat{\alpha}_0 \geq 0$ e $\hat{v}_k^{(1)}$ tais que $\hat{v}_k^{(1)} = \hat{v}_k^{(0)} + \hat{\alpha}_0 \Delta \hat{v}_k^{(0)}$ é viável;
 - 15: Resolver o problema convexo (3.32) para \hat{v}_k considerando $\hat{v}_k^{(1)}$ como ponto inicial;
 - 16: $d_k = \hat{v}_k - v_k$; (direção de descida global)
 - 17: **saída** d_k ;
 - 18: **fim**
-

Para a direção s_k , a direção de curvatura negativa $s_k^{(1)}$ do Algoritmo 1 é re-escalada. Primeiro definimos

$$s_k^{(2)} = \begin{cases} -s_k^{(1)} & , \text{ se } \nabla M(v_k; y_k^E, \mu_k^R)^T s_k^{(1)} > 0 \\ s_k^{(1)} & , \text{ caso contrário.} \end{cases}$$

Isto implica que $s_k^{(2)}$ é uma direção não crescente para $M(v_k; y_k^E, \mu_k^R)$ em v_k . É feito o re-escalamento $s_k = \sigma_k s_k^{(2)}$ para que $v_k + s_k + d_k$ seja viável e a parte primal de s_k e d_k tenham normas comparáveis.

Algoritmo 3 (direção viável de curvatura negativa s_k para a função de mérito)

- 1: **Dados** $s_k^{(1)}, \xi_k, x_k, d_k, J_k, H_k$
 - 2: $s_k^{(2)} = \begin{cases} -s_k^{(1)} & , \text{ se } \nabla M(v_k, y_k^E, \mu_k^R)^T s_k^{(1)} > 0; \\ s_k^{(1)} & , \text{ caso contrário;} \end{cases}$
 - 3: Considerar p_k e $u_k^{(2)}$ as primeiras n componentes de d_k e $s_k^{(2)}$;
 - 4: $\sigma_k = \arg \max_{\sigma \geq 0} \left\{ \sigma : x_k + p_k + \sigma u_k^{(2)} \geq 0, \left\| \sigma u_k^{(2)} \right\| \leq \max(\xi_k, \|p_k\|) \right\}$;
 - 5: $s_k = \sigma_k s_k^{(2)}$;
 - 6: **saída** s_k ;
 - 7: **fim**
-

3.3.4 Busca linear flexível

A ideia de uma busca linear flexível foi proposta por Curtis e Nocedal [15]. Seja $\Delta v_k = d_k + s_k$. Se $d_k \neq 0$, $s_k = 0$ e (x_k, y_k) provém de uma V - O iteração então consideramos α_k satisfazendo

$$M(v_k + \alpha_k \Delta v_k, y_k^E, \mu) \leq M(v_k, y_k^E, \mu) + \gamma_S \alpha_k \nabla M(v_k, y_k^E, \mu)^T \Delta v_k.$$

Caso contrário, consideramos

$$\begin{aligned} M(v_k + \alpha_k \Delta v_k, y_k^E, \mu) &\leq M(v_k, y_k^E, \mu) + \gamma_S \alpha_k \nabla M(v_k, y_k^E, \mu)^T \Delta v_k \\ &\quad + \frac{1}{2} \alpha_k^2 \min(0, (\Delta v_k)^T B(v_k; \mu_{k-1}^R) \Delta v_k) \end{aligned}$$

onde $0 < \gamma_S < 1$.

Na condições anteriores precisamos que as desigualdades anteriores sejam válidas para $\mu \in [\mu_k^R, \mu_k]$, na prática testamos até que se consiga para $\mu = \mu_k^R$ ou $\mu = \mu_k$. Para juntar ambas condições em uma, definimos a função

$$\Psi_k(\alpha, \mu) = M(v_k + \alpha \Delta v_k, y_k^E, \mu)$$

com y_k^E fixo e a função

$$\psi_k(\alpha, \mu, \ell_k) = \Psi_k(0, \mu) + \alpha \Psi'_k(0, \mu) + \frac{1}{2}(\ell_k - 1)\alpha^2 \min(0, (\Delta v_k)^T B(v_k; \mu_{k-1}^R) \Delta v_k), \quad (3.36)$$

onde Ψ'_k denota a derivada com respeito a α e ℓ assume os valores 1 ou 2 dependendo da ordem requerida no modelo de busca linear. No algoritmo se requer que o passo α_k produza uma redução expressa em termos da condição de Armijo, como

$$\Psi_k(0, \mu) - \Psi_k(\alpha_k, \mu) \geq \gamma_S (\psi_k(0, \mu_k^R, \ell_k) - \psi_k(\alpha_k, \mu_k^R, \ell_k)) \quad (3.37)$$

e na prática o passo é encontrado reduzindo α_k por um valor constante até que se tenha $\rho_k(\alpha_k, \mu_k, \ell_k) \geq \gamma_S$ ou $\rho_k(\alpha_k, \mu_k^R, \ell_k) \geq \gamma_S$, onde

$$\rho_k(\alpha, \mu, \ell_k) = \frac{\Psi_k(0, \mu) - \Psi_k(\alpha, \mu)}{\psi_k(0, \mu_k^R, \ell_k) - \psi_k(\alpha, \mu_k^R, \ell_k)}.$$

Algoritmo 4 (busca linear flexível)

- 1: **Dados** $d_k, s_k, y_k^E, \mu_k, \mu_k^R, \mu_{k-1}^R, \ell_k, J_k, H_k$
 - 2: **Constantes:** $\gamma_S \in (0, \frac{1}{2})$;
 - 3: Computar $\nabla M = \nabla M(x_k, y_k, y_k^E, \mu_k^R)$;
 - 4: **se** $s_k = 0$ **e** $d_k = 0$ **então**
 - 5: $\alpha_k = 1$;
 - 6: **caso contrário** ($d_k \neq 0$ **ou** $\nabla M^T s_k < 0$ **ou** $\mu_k^R = \mu_{k-1}^R$) **então**
 - 7: $\alpha_k = 1$;
 - 8: **enquanto** $\rho_k(\alpha_k; \mu_k^R, \ell_k) < \gamma_S$ **e** $\rho_k(\alpha_k; \mu_k, \ell_k) < \gamma_S$
 - 9: $\alpha_k = \frac{1}{2}\alpha_k$;
 - 10: **fim**
 - 11: **caso contrário** [$d_k = 0, s_k \neq 0, \xi_k > 0$]
 - 12: $\xi_k^R = -s_k^T \nabla^2 M(x_k, y_k, y_k^E, \mu_k^R) s_k / \|u_k\|^2$; [Por definição, $s_k = (u_k, w_k)$]
 - 13: **se** $\xi_k^R > \gamma_S \xi_k$ **então**
 - 14: $\alpha_k = 1$;
 - 15: **enquanto** $\rho_k(\alpha_k, \mu_k^R, \ell_k) < \gamma_S$ **e** $\rho_k(\alpha_k, \mu_k, \ell_k) < \gamma_S$
 - 16: $\alpha_k = \frac{1}{2}\alpha_k$;
 - 17: **fim**
 - 18: **caso contrário**
 - 19: $\alpha_k = 0$;
 - 20: **fim**
 - 21: **fim**
 - 22: **saída** $\alpha_k \geq 0$
 - 23: **fim**
-

Finalmente atualizamos o parâmetro de penalização

$$\mu_{k+1} = \begin{cases} \mu_k & \text{se } \rho_k(\alpha_k; \mu_k) \geq \gamma_s, \text{ ou } d_k = s_k = 0, \text{ ou } \alpha_k = 0 \\ \max\left(\frac{1}{2}\mu_k, \mu_k^R\right) & \text{no outro caso} \end{cases} \quad (3.38)$$

e definimos

$$v_{k+1} = v_k + \alpha_k(d_k + s_k).$$

3.3.5 Critério de parada

O algoritmo é formalmente estabelecido no Algoritmo 5 e possui dois critérios de parada. O primeiro é

$$r(x_k, y_k) \leq \tau_s, \quad \xi_k \leq \tau_s \text{ e } \mu_{k-1}^R \leq \tau_s \quad (3.39)$$

onde τ_s é uma tolerância de parada positiva, $r(x, y)$ é definido em (3.2) e ξ_k é computado no Algoritmo 1. Quando as condições (3.39) são satisfeitas, (x_k, y_k) é um ponto KKT aproximado de segunda ordem para o Problema (3.1). O segundo, consiste em identificar convergência a um ponto estacionário não viável, isto é um minimizador do problema

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & \frac{1}{2} \|c(x)\|_2^2 \\ \text{s.a.} \quad & x \geq 0 \end{aligned} \quad (3.40)$$

no qual a restrição $c(x)$ não é válida. As condições de otimalidade de primeira ordem para o Problema (3.40) são dadas por

$$\begin{aligned} J(x)^T c(x) - z &= 0 \\ z_i &\geq 0 \\ x_i z_i &= 0 \quad \forall i = 1, \dots, n \end{aligned}$$

que equivale a

$$\|\min(x, J(x)^T c(x))\| = 0$$

o que motiva o segundo critério de parada

$$\min(\|c(x_k)\|, \tau_s) > \mu_k^R, \quad \|\min(x_k, J(x_k)^T c(x_k))\| \leq \tau_s \text{ sendo } k \text{ uma M-iteração.} \quad (3.41)$$

Algoritmo 5 (PQS primal-dual de segunda ordem)

- 1: **Dados** x_1, y_1, μ_0^R, μ_1
- 2: **Constantes:** $\{\tau_s, \gamma_s\} \subset (0, \frac{1}{2})$, $0 < \bar{\gamma} < 1$, $y_{\max} > 0$ e $0 < \epsilon_a \ll 1$;
- 3: escolher $y_0^E \in \mathbb{R}^m$, $\{\tau_0, \phi_{V,0}^{\max}, \phi_{O,0}^{\max}\} \subset (0, \infty)$ e $0 < \mu_0^R \leq \mu_1 < \infty$;
- 4: $k = 1$;

```

5:   repita
6:     Computar o conjunto  $\epsilon$ -livre  $\mathcal{F}_\epsilon(x_k, y_k, \mu_{k-1}^R)$  de (3.18)
7:      $J_k = J(x_k)$ ;  $H_k = H(x_k, y_k)$ ;
8:      $(s_k^{(1)}, \xi_k) = \text{menor\_curvatura}(x_k, y_k, \mu_{k-1}^R, J_k, H_k)$ ; [Algoritmo 1]
9:     Calcular  $r(x_k, y_k)$  de (3.2);
10:    se (vale condição de parada (3.39)) então
11:      saída o ponto KKT de segunda ordem  $(x_k, y_k)$ ; (Definição 1.14);
12:    fim
13:    se  $(\phi_V(x_k, y_k) \leq \frac{1}{2}\phi_{V,k-1}^{\max})$  e  $\|y_k\| \leq y_{\max}$  então [V-iteração]
14:       $\phi_{V,k}^{\max} = \frac{1}{2}\phi_{V,k-1}^{\max}$ ;  $y_k^E = y_k$ ;  $\tau_k = \tau_{k-1}$ ;
15:      Atualizar  $\mu_k^R$  como em (3.26);
16:    se em outro caso  $(\phi_O(x_k, y_k) \leq \frac{1}{2}\phi_{O,k-1}^{\max})$  e  $\|y_k\| \leq y_{\max}$  [O-iteração]
17:       $\phi_{O,k}^{\max} = \frac{1}{2}\phi_{O,k-1}^{\max}$ ;  $y_k^E = y_k$ ;  $\tau_k = \tau_{k-1}$ ;
18:      Atualizar  $\mu_k^R$  como em (3.26);
19:    se em outro caso  $((x_k, y_k)$  satisfaz (3.28)) então [M-iteração]
20:      Atualizar  $y_k^E$  como em (3.29);  $\tau_k = \frac{1}{2}\tau_{k-1}$ ;
21:      Atualizar  $\mu_k^R$  como em (3.30);
22:    caso contrário
23:       $y_k^E = y_{k-1}^E$ ;  $\tau_k = \tau_{k-1}$ ;  $\mu_k^R = \mu_{k-1}^R$ ; [F-iteração]
24:    fim
25:    se (vale condição de parada (3.41)) então
26:      saída com o ponto estacionário aproximado não viável  $x_k$ ;
27:    fim
28:    Computar o conjunto  $\epsilon$ -livre  $\mathcal{F}_\epsilon(x_k, y_k, \mu_k^R)$  de (3.18);
29:     $d_k = \text{direção\_descida}(x_k, y_k, \mu_k^R, J_k, H_k)$ ; [Algoritmo 2]
30:     $s_k = \text{direção\_curvatura\_negativa}(s_k^{(1)}, \xi_k, x_k, d_k, J_k, H_k)$ ; [Algoritmo 3]
31:    se  $(d_k \neq 0$  e  $s_k = 0$  e  $(x_k, y_k)$  é uma V-O iteração) então
32:       $\ell_k = 1$ ;
33:    caso contrário
34:       $\ell_k = 2$ ;
35:    fim
36:     $\mu_k = \max(\mu_k^R, \mu_k)$ ;
37:     $\alpha_k = \text{busca\_linear\_flexível}(d_k, s_k, y_k^E, \mu_k, \mu_k^R, \mu_{k-1}^R, \ell_k, J_k, H_k)$ ; [Algoritmo 4]
38:    atualizar  $\mu_{k+1}$  como em (3.38);
39:     $v_{k+1} = (x_{k+1}, y_{k+1}) = v_k + \alpha_k d_k + \alpha_k s_k$ ;
40:     $k = k + 1$ ;
41:  fim
42: fim

```

3.4 Propriedades do Algoritmo 5

Para a análise do algoritmo é necessário fazer algumas considerações que garantam a convergência. Neste sentido, assumiremos que as funções f e c são suficientemente diferenciáveis, que as Hessianas aproximadas \hat{H} são limitadas e que existe um ponto limite da sequência gerada $\{x_k\}$. Além disso, os critérios de parada serão omitidos. Essas considerações nos permitem estudar as propriedades associadas a um número infinito de iterações. Consideraremos as mesmas condições de [30] que são as seguintes:

Condição 3.6. *Considerando a sequência $\{(x_k, y_k)\}$ gerada pelo Algoritmo 5, vamos assumir que:*

(a) *A sequência de matrizes $\{\hat{H}(x_k, y_k)\}_{k \geq 0}$ é escolhida de tal maneira que*

$$\left\| \hat{H}(x_k, y_k) \right\| \leq \hat{H}_{\max} \text{ e } \lambda_{\min}(\hat{H}(x_k, y_k) + \frac{1}{\mu_{k-1}^R} J(x_k)^T J(x_k)) \geq \lambda_{\min}$$

para \hat{H}_{\max} e λ_{\min} positivos e todo $k \geq 0$.

(b) *As funções f e c são duas vezes continuamente diferenciáveis.*

(c) *A sequência $\{x_k\}_{k \geq 0}$ possui uma subsequência convergente.*

(d) *Os critérios de parada das linhas 10 e 25 são omitidos do Algoritmo 5.*

Pelo comentado na fundamentação do método PQSe (Seção 3.2), nos interessa que a atualização do multiplicador estimado y_k^E seja se possível igual a y_k . Nos seguintes resultados veremos que de fato isto acontece no algoritmo. Primeiramente veremos algumas propriedades que acontecem se o conjunto de iterações de tipo O, V e F é finito.

Lema 3.7. *[30, Teorema 3.5] Considerando a sequência infinita $\{v_k\}$ onde $v_k = (x_k, y_k)$ é gerada pelo Algoritmo 5. Suponha que existe \hat{k} tal que a k -ésima iteração é do tipo F para todo $k \in S = \{k \mid k \geq \hat{k}\}$, então*

(1) *Existem constantes positivas τ , μ^R e μ , e um vetor constante y^E tal que*

$$\tau_k = \tau, \quad \mu_k^R = \mu^R, \quad \mu_k = \mu \quad \text{e} \quad y_k^E = y^E \text{ para todo } k \in S. \quad (3.42)$$

(2) *As sequências $\{d_k\}$, $\{y_k\}$ e $\{\pi(x_k, y^E, \mu^R)\}$ ($\pi(x, y^E, \mu) = y^E - c(x)/\mu$) são uniformemente limitadas em S .*

(3) *As sequências $\{\xi_k\}$, $\{s_k\}$, $\{\Delta v_k\}$, $\{B(v_k; \mu^R)\}$, $\{\nabla^2 M(v_k; y^E, \mu^R)\}$, $\{\Delta v_k^T B(v_k; \mu^R) \Delta v_k\}$ e $\{\Delta v_k^T \nabla^2 M(v_k; y^E, \mu^R) \Delta v_k\}$ são uniformemente limitadas em S .*

(4) Se $\nabla M(v_k; y^E, \mu^R)^T d_k \xrightarrow{S'} 0$ para toda subsequência $S' \subseteq S$ então

$$d_k \xrightarrow{S'} 0, \quad \|\pi(x_k, y^E, \mu^R) - y_k\| \xrightarrow{S'} 0, \quad \|B(v_k; \mu^R) - \nabla^2 M(v_k; y^E, \mu^R)\| \xrightarrow{S'} 0$$

e $\xi_k > \tau$ para todo $k \in S'$ suficientemente grande.

(5) Existe $\epsilon_F > 0$ tal que

$$\nabla M(v_k; y^E, \mu^R)^T d_k \leq -\epsilon_F \quad \text{ou} \quad s_k^T B(v_k; \mu^R) s_k \leq -\epsilon_F \quad \text{para todo } k \in S.$$

Demonstração. Como a definição dos iterandos do tipo F é igual à considerada no Algoritmo 5 em [30], então a prova é a mesma que a prova do [30, Teorema 3.5]. \square

O resultado anterior permite garantir que o número de iterandos de tipo O, V e M é infinito, como veremos a seguir. A demonstração segue os passos da feita em [30], porém, considerando que no nosso caso os iterandos de tipo O e V exigem uma condição a mais para ser considerados como tal (y_k^E limitado), veremos a prova em detalhe.

Teorema 3.8. *Considerando a sequência infinita $\{v_k\}$ onde $v_k = (x_k, y_k)$ é gerada pelo Algoritmo 5:*

(1) O número de iterandos do tipo O, V ou M é infinito.

(2) A sequência de parâmetros de regularização satisfaz $\mu_k^R \rightarrow 0$.

Demonstração. A prova do item (1) será feita por contradição. Suponhamos que o número de iterandos do tipo O, V e M é finito. Então existe um índice \hat{k} tal que se $k > \hat{k}$ então todos os iterandos são de tipo F. Pelo Lema 3.7 (1) temos que (3.42) é válida no conjunto $S = \{k \mid k \geq \hat{k}\}$. Será mostrado que este resultado implica a existência de uma subsequência infinita $S_1 \subseteq S$ e de valores positivos fixos μ^R e κ tais que

$$M(v_{k+1}; y^E, \mu^R) \leq M(v_k; y^E, \mu^R) - \kappa \quad \text{para todo } k \in S_1 \quad (3.43)$$

o que implica que $M(v_k; y^E, \mu^R) \rightarrow -\infty$. Porém isto é impossível pois pelas Condições 3.6 e a parte (2) do Lema 3.7, M é limitada em S_1 .

Se cada iteração é de tipo F no conjunto S , então o Lema 3.7 (1) implica que $\mu_k^R = \mu^R$, o Algoritmo 5 implica $\ell_k = 2$, e o Algoritmo 2 gera uma direção d_k de descida global para todo $k \in S$. Considerando $\mu = \mu^R$ e $\ell_k = 2$ em (3.37) temos

$$\Psi_k(\alpha, \mu^R) \leq \Psi_k(0, \mu^R) + \gamma_S (\psi_k(\alpha, \mu_k^R, 2) - \psi_k(0, \mu_k^R, 2))$$

onde $\Psi_k(\alpha, \mu) = M(v_k + \alpha \Delta v_k, y_k^E, \mu)$, $\Delta v_k = d_k + s_k$ e $\psi_k(\alpha, \mu_k^R, 2)$ é dado pela equação

(3.36), isto é,

$$\psi_k(\alpha, \mu, 2) = \Psi_k(0, \mu) + \alpha \Psi'_k(0, \mu) + \frac{1}{2} \alpha^2 \min(0, (\Delta v_k)^T B(v_k; \mu_{k-1}^R) \Delta v_k).$$

Note que

$$\psi_k''(\alpha, \mu, \ell_k) = \min(0, (\Delta v_k)^T B(v_k; \mu_{k-1}^R) \Delta v_k) \leq 0. \quad (3.44)$$

Logo o critério de Armijo está dado por

$$\Psi_k(\alpha, \mu^R) \leq \Psi_k(0, \mu^R) + \gamma_S \alpha \Psi'_k(0, \mu^R) + \frac{1}{2} \alpha^2 \gamma_S \psi_k''(\alpha, \mu, \ell_k). \quad (3.45)$$

Usando expansão de Taylor em (3.45), temos

$$\alpha \Psi'_k(0, \mu^R) + \frac{\alpha^2}{2} \Psi_k''(0, \mu^R) + O(|\alpha|^3) \leq \gamma_S \alpha \Psi'_k(0, \mu^R) + \frac{1}{2} \alpha^2 \gamma_S \psi_k''(\alpha, \mu, \ell_k),$$

de onde segue que

$$(1 - \gamma_S) \Psi'_k(0, \mu^R) + \frac{\alpha}{2} (\Psi_k''(0, \mu^R) - \gamma_S \psi_k''(\alpha, \mu^R, \ell_k)) \leq -O(|\alpha|^2). \quad (3.46)$$

Dois casos são considerados. Primeiramente seja $\Psi'_k(0, \mu^R) = \nabla M(v_k, y_k^E, \mu^R)^T \Delta v_k \leq -\delta$ para algum $\delta > 0$ e k suficientemente grande. Note que pela definição de Ψ_k , $\Psi_k''(0, \mu^R) = \Delta v_k^T \nabla^2 M(v_k; y^E, \mu^R) \Delta v_k$, logo pelas partes (2) e (3) do Lema 3.7 $\{\Psi_k''(0, \mu^R)\}_{k \in S}$ e $\{\psi_k''(0, \mu^R)\}_{k \in S}$ são uniformemente limitadas. Como neste caso $\Psi'_k(0, \mu^R) \leq -\delta$ então existe $\hat{\alpha}$ tais que (3.46) é satisfeito para todo $0 < \alpha < \hat{\alpha}$ e $k \in S_1$ suficientemente grande. Logo a busca linear terminaria com um valor $\alpha_k \geq \alpha_{\min}$ para algum $\alpha_{\min} > 0$ para todo k suficientemente grande. Pelo Lema 3.7 (1) temos de (3.45) e (3.44) que

$$\Psi_k(\alpha_k, \mu^R) \leq \Psi_k(0, \mu^R) - \gamma_S \delta \alpha_{\min} \quad \text{para todo } k \text{ suficientemente grande.}$$

Denotando por S_1 o conjunto de k suficientemente grande tal que $\Psi'_k(0, \mu^R) \leq -\delta$, temos

$$M(v_{k+1}; y^E, \mu^R) = M(v_k + \alpha \Delta v_k, y_k^E, \mu^R) \leq M(v_k; y^E, \mu^R) - \gamma_S \delta \alpha_{\min} \quad \forall k \in S_1.$$

Num segundo caso considere que exista uma sequência S_1 , tal que $\Psi'_k(0, \mu^R) \xrightarrow{S_1} 0$. O Algoritmo 5 encontra direções d_k e s_k tais que

$$\nabla M(v_k; y^E, \mu^R)^T d_k \leq 0 \quad \text{e} \quad \nabla M(v_k; y^E, \mu^R)^T s_k \leq 0. \quad (3.47)$$

Logo

$$\Psi'_k(0, \mu^R) = \nabla M(v_k; y^E, \mu^R)^T \Delta v_k = \nabla M(v_k; y^E, \mu^R)^T d_k + \nabla M(v_k; y^E, \mu^R)^T s_k \leq 0. \quad (3.48)$$

Assim para que a condição de Armijo (3.46) seja satisfeita é suficiente que

$$\frac{1}{2} (\Psi_k''(0, \mu^R) - \gamma_S \psi_k''(\alpha, \mu, \ell_k)) \leq -O(|\alpha|). \quad (3.49)$$

Além disso por (3.47) e (3.48)

$$\nabla M(v_k; y^E, \mu^R)^T d_k \xrightarrow{S_1} 0. \quad (3.50)$$

Pelo Lema 3.7 (5), existe $\epsilon_F > 0$ tal que

$$s_k^T B(v_k; \mu^R) s_k \leq -\epsilon_F \text{ para todo } k \in S_1 \text{ suficientemente grande.} \quad (3.51)$$

Como

$$\begin{aligned} \Delta v_k^T B(v_k; \mu^R) \Delta v_k &= (d_k + s_k)^T B(v_k; \mu^R) (d_k + s_k) \\ &= d_k^T B(v_k; \mu^R) d_k + d_k^T B(v_k; \mu^R) s_k + s_k^T B(v_k; \mu^R) d_k + s_k^T B(v_k; \mu^R) s_k, \end{aligned}$$

pelos partes (3) e (4) do Lema 3.7 junto com (3.50), os três primeiros termos da soma anterior convergem para zero. Levando em conta a última afirmação, (3.51) e a definição de ψ

$$\psi_k''(\alpha, \mu, \ell_k) = \Delta v_k^T B(v_k; \mu^R) \Delta v_k \leq -\frac{1}{2} \epsilon_F \text{ para todo } k \in S_1 \text{ suficientemente grande.} \quad (3.52)$$

Considerando a definição de Ψ_k e as partes (3) e (5) do Lema 3.7 e (3.52), temos que para todo $k \in S_1$ suficientemente grande

$$\|\Psi_k''(0, \mu^R) - \psi_k''(\alpha, \mu, \ell_k)\| = \|\Delta v_k^T \nabla^2 M(v_k; y^E, \mu^R) \Delta v_k - \Delta v_k^T B(v_k; \mu^R) \Delta v_k\| \xrightarrow{S_1} 0.$$

Logo por (3.52)

$$\begin{aligned} \Psi_k''(0, \mu^R) - \gamma_S \psi_k''(\alpha, \mu, \ell_k) &= \Psi_k''(0, \mu^R) - \psi_k''(\alpha, \mu, \ell_k) + (1 - \gamma_S) \psi_k''(\alpha, \mu, \ell_k) \\ &\leq \frac{1}{2} (1 - \gamma_S) \Delta v_k^T B(v_k; \mu^R) \Delta v_k \\ &\leq \frac{1}{4} (1 - \gamma_S) \epsilon_F \text{ para } k \in S_1 \text{ suficientemente grande.} \end{aligned}$$

Por (3.49), de esta última desigualdade é suficiente considerar:

$$-\frac{1}{8} (1 - \gamma_S) \epsilon_F \leq -O(|\alpha|).$$

para que a condição de Armijo seja satisfeita. Assim, existe $\hat{\alpha}$ tal que (3.46) é satisfeito para todo $0 < \alpha < \hat{\alpha}$ e $k \in S_1$ suficientemente grande. Logo, como no outro caso a busca linear termina com um valor $\alpha_k \geq \alpha_{\min}$ para algum $\alpha_{\min} > 0$ para todo k suficientemente

grande. Logo em (3.45), considerando (3.52) e o fato que neste caso $\Psi'_k(0, \mu^R) \xrightarrow{S_1} 0$, temos

$$\begin{aligned}\Psi_k(\alpha_k, \mu^R) &\leq \Psi_k(0, \mu^R) + \gamma_S \alpha_k \Psi'_k(0, \mu^R) - \frac{1}{2} \alpha_{\min}^2 \gamma_S \psi''_k(\alpha_k, \mu, \ell_k) \\ &\leq \Psi_k(0, \mu^R) - \frac{1}{4} \alpha_{\min}^2 \epsilon_F \text{ para todo } k \in S_1 \text{ suficientemente grande.}\end{aligned}$$

Logo

$$M(v_{k+1}, y^E, \mu^R) = M(v_k + \alpha \Delta v_k, y_k^E, \mu^R) \leq M(v_k, y^E, \mu^R) - \frac{1}{4} \alpha_{\min}^2 \epsilon_F \quad \forall k \in S_1.$$

Para a demonstração de item (2), temos por (3.30) que no caso de iterações do tipo M que $\mu_k^R \leq \frac{1}{2} \mu_{k-1}^R$ e no caso de iterações do tipo F que $\mu_k^R = \mu_{k-1}^R$. Se o número de iterações do tipo O e V for finito, então pela parte (1) o número de iterações de tipo M é infinito logo $\mu_k^R \rightarrow 0$. Assumimos então que o número de iterações de tipo O e V é infinito. Seja S_2 o conjunto de índices dos iterandos do tipo V. Sem perda de generalidade, assumimos que S_2 é infinito, então por, (3.24), $\phi_V(v_k, \xi_k) \leq \frac{1}{2} \phi_{V,k-1}^{\max}$ para todo $k \in S_2$. Como por, (3.25), $\phi_{V,k}^{\max} = \frac{1}{2} \phi_{V,k-1}^{\max}$, então $\phi_V(v_k, \xi_k) \xrightarrow{S_2} 0$. Pela definição ϕ_V em (3.22) temos que

$$c(x_k) \xrightarrow{S_2} 0, \quad \|\min(x_k, \nabla f(x_k) - J(x_k)^T y_k)\| \xrightarrow{S_2} 0 \quad \text{e} \quad \xi_k \xrightarrow{S_2} 0$$

e isto implica

$$r(x_k, y_k) = \|(c(x_k), \min(x_k, \nabla f(x_k) - J(x_k)^T y_k))\| \xrightarrow{S_2} 0,$$

de onde pela atualização de μ_k^R em (3.26) $\mu_k^R \rightarrow 0$. □

Neste ponto devemos observar que a modificação feita no Algoritmo 5 de [30] possivelmente gera sequências diferentes devido a que se pede uma condição a mais para ser uma iteração do tipo O e V (y_k^E limitado), porém esta condição não atrapalha na prova do teorema, posto que propriedades que possuem os iterandos não se modificam. Isto também é observado nos outros resultados da análise de convergência feitos em [30, 31]. Assim, consideramos esses resultados válidos para nosso algoritmo, e no que se segue mencionamos os que são de nosso interesse.

Teorema 3.9. [30, Teorema 3.7] *Seja \mathcal{M} o conjunto formado pelos índices dos iterandos de tipo M. Se o conjunto de iterandos de tipo O e V é finito, então S é infinito e todo ponto limite x_* de $\{x_k\}_{k \in \mathcal{M}}$ satisfaz $c(x_*) \neq 0$ e é um ponto KKT do problema*

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & \frac{1}{2} \|c(x)\|^2 \\ \text{s.a.} \quad & x \geq 0. \end{aligned}$$

Como nosso interesse é que $y_k^E = y_k$ sempre que possível, e isto acontece quando se tem iterações dos tipos O e V, o último resultado, junto com o Teorema 3.8 (1), mostra que para analisar a convergência é necessário assumir que o conjunto

$$\mathcal{S} = \{k \mid (x_k, y_k) \text{ é uma V-O iteração}\} \quad (3.53)$$

é infinito. O Lema a seguir resume outros resultados de nosso interesse que estão associados ao conjunto \mathcal{S} . Para isso, como foi feito em [31], será necessário impor as seguintes condições.

Condição 3.10. [31, Hipóteses 2 e 3] *Considerando a sequência $\{(x_k, y_k)\}$ gerada pelo Algoritmo 5. Assumimos que:*

- (a) *O conjunto \mathcal{S} é infinito e existe uma subsequência $\mathcal{S}_* \subseteq \mathcal{S}$ tal que $\lim_{k \in \mathcal{S}_*} (x_k, y_k) \rightarrow (x^*, y^*)$ onde (x^*, y^*) é um par KKT de primeira ordem para o Problema (3.1).*
- (b) *Se (x^*, y^*) é o par KKT de primeiro ordem de (3.1) e $\mathcal{Y}(x^*) = \{y \in \mathbb{R}^m \mid (x^*, y) \text{ satisfaz } r(x^*, y) = 0\}$, então*
 - 1. *existe um conjunto compacto $\Lambda(x^*) \subseteq \mathcal{Y}(x^*)$ tal que y^* pertence ao interior (não vazio) de $\Lambda(x^*)$ relativo a $\mathcal{Y}(x^*)$ e*
 - 2. *(x^*, y) cumpre a condição suficiente de segunda ordem (Definição 1.14) para todo $y \in \Lambda(x^*)$.*

Lema 3.11. [31, Lemas 1,2,3,4,6 e 7; Teoremas 2 e 3] *Para $\epsilon > 0$ suficientemente pequeno e ρ tal que $1 < \rho < 1 + \bar{\gamma}$ ($0 < \bar{\gamma} < 1$) existe uma V-iteração e um índice $k_V = k_V(\epsilon)$ tal que para todo $k \geq k_V$*

- (a) *Os conjuntos ϵ -ativos satisfazem $\mathcal{A}_\epsilon(x_k, y_k, \mu_{k-1}^R) = \mathcal{A}_\epsilon(x_k, y_k, \mu_k^R) = \mathcal{A}(x^*)$ e os conjuntos ϵ -livres satisfazem $\mathcal{F}_\epsilon(x_k, y_k, \mu_{k-1}^R) = \mathcal{F}_\epsilon(x_k, y_k, \mu_k^R) = \mathcal{F}(x^*)$.*
- (b) *A direção viável de curvatura negativa s_k é zero e a busca linear encontra $\alpha_k = 1$.*
- (c) *$\mu_k^R = r(x_k, y_k)^{\bar{\gamma}} > 0$.*
- (d) *A direção de descida local $d_k = (p_k, q_k)$ associada ao Problema (3.34) é computada e todas as condições de (3.35) são satisfeitas. Além disso*

$$\max\{\|p_k\|, \|q_k\|\} \leq O(r(x_k, y_k)).$$

- (e) *$(\hat{x}_k, \hat{y}_k) = (x_k + p_k, y_k + q_k)$ satisfaz*

$$\|\hat{x}_k - x^*\| \leq O(r(x_k, y_k)^{1+\bar{\gamma}}).$$

Outro resultado importante está relacionado à ordem de convergência, e como é de se esperar no caso PQS tem-se o seguinte resultado:

Teorema 3.12. *[31, Teorema 4] Se $\lim_{k \rightarrow \infty} (x_k, y_k) \rightarrow (x^*, y^*)$, a convergência é superlinear.*

O objetivo deste capítulo foi descrever um algoritmo baseado em Programação Quadrática Sequencial de segunda ordem projetado para resolver problemas de otimização não linear, o qual possui ordem de convergência superlinear. Como foi visto nos Capítulos anteriores, os problemas MPEC não cumprem algumas das condições de qualificação usuais, e seus minimizadores em geral não satisfazem as condições KKT. Motivo pelo qual, no Capítulo 2 foram introduzidas novas condições de qualificação e novos conceitos de estacionariedade. No capítulo seguinte, mostraremos que as novas condições de qualificação garantem que pontos de acumulação viáveis, de sequências geradas pelo Algoritmo 5 para quando aplicado a problemas MPEC, satisfazem estes novos conceitos de estacionariedade.

Capítulo 4

Convergência para pontos S-estacionários e M-estacionários

Neste capítulo apresentamos os principais resultados deste trabalho de tese, os quais estão relacionados com aplicar o Algoritmo 5 a problemas MPEC do tipo

$$\begin{aligned} \min \quad & f(x) \\ \text{s.a} \quad & G(x) \geq 0, \quad H(x) \geq 0, \\ & G(x)^T H(x) \leq 0 \\ & x \geq 0 \end{aligned} \tag{4.1}$$

$f : \mathbb{R}^n \rightarrow \mathbb{R}$, $G, H : \mathbb{R}^n \rightarrow \mathbb{R}^l$. Note que neste caso a restrição de complementaridade é de desigualdade, porém esta representação é equivalente ao caso com igualdade posto que $G(x) \geq 0$ e $H(x) \geq 0$. A vantagem de considerar com desigualdade é que os multiplicadores associados à restrição de complementaridade ganham sinal. A análise é feita sobre pontos de acumulação viáveis da sequência gerada pelo Algoritmo 5 e apresentamos dois resultados. No primeiro, considerando a condição MPEC-RCLPD e uma propriedade de limitação vinculada ao multiplicador da restrição de complementaridade, mostraremos que é possível garantir convergência a pontos S-estacionários. No segundo consideramos a condição MPEC-LICQ, garantido neste caso convergência para pontos que são pelo menos M-estacionários. Finalmente apresentamos os resultados dos testes numéricos considerando problemas da coleção MacMPEC (<http://www.mcs.anl.gov/~leyffer/MacMPEC>) para validar nossos resultados.

Para analisar os pontos de acumulação gerados pelo Algoritmo 5, encontraremos uma maneira explícita para a sequência de multiplicadores aproximados associada a estes pontos, o que permitirá fazer uma análise de convergência. Posto que o Algoritmo 5 está projetado para lidar com problemas com restrições de igualdade, analisaremos algumas propriedades quando o algoritmo é aplicado a problemas com desigualdade, dos quais o Problema MPEC (4.1) é um caso particular.

4.1 Pontos Aproximadamente KKT

Para poder utilizar o Algoritmo 5, nosso primeiro objetivo é mostrar que este gera pontos AKKT. Esta análise permitirá fixar algumas sequências que serão úteis na análise de convergência quando o Algoritmo 5 for aplicado a problemas MPEC. Em [62] é demonstrado que isto de fato acontece, porém para um melhor entendimento veremos uma prova detalhada. Consideraremos a definição de pontos AKKT [4] adaptada ao Problema (3.1)

$$\begin{aligned} \min \quad & f(x) \\ \text{s.a} \quad & c(x) = 0 \\ & x \geq 0 \end{aligned}$$

onde $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $c : \mathbb{R}^n \rightarrow \mathbb{R}^p$ são funções continuamente diferenciáveis. Dizemos que $x^* \in \mathbb{R}^n$ é um ponto AKKT se existem sequências $\{x_k\} \in \mathbb{R}^n$, $\{y_k\} \in \mathbb{R}^p$, $\{w_k\} \in \mathbb{R}^n$ e $\{\varepsilon_k\} \in \mathbb{R}_+$ tais que $x_k \rightarrow x^*$, $\varepsilon_k \rightarrow 0^+$ e:

$$w_k \geq 0, \quad (4.2a)$$

$$\|c(x_k)\| \leq \varepsilon_k, \quad (4.2b)$$

$$\|\max(0, -x_k)\| \leq \varepsilon_k, \quad (4.2c)$$

$$\|\nabla f(x_k) - J(x_k)^T y_k - w_k\| \leq \varepsilon_k, \quad (4.2d)$$

$$[w_k]_i = 0 \text{ se } [x_k]_i > \varepsilon_k. \quad (4.2e)$$

Proposição 4.1. *O Algoritmo 5 gera pontos AKKT.*

Demonstração. Consideremos que o conjunto \mathcal{S} , dado por (3.53), é infinito. Se $x^* \in \mathbb{R}^n$ é um ponto de acumulação da sequência $\{x_k\}_{\mathcal{S}}$ então afirmamos que x^* é um ponto AKKT. De fato, sem perda de generalidade suponhamos que $x_k \rightarrow x^*$, passando a uma subsequência se necessário. Como (x_k, y_k) é uma V-O iteração, pela atualização de $\phi_{V,k}^{\max}$ e $\phi_{O,k}^{\max}$ (3.25) temos por (3.24) que $\phi_V(x_k, y_k, \xi_k) \rightarrow 0$ ou $\phi_O(x_k, y_k, \xi_k) \rightarrow 0$, o que implica de (3.21) e (3.22) que $\|c(x_k)\| \rightarrow 0$ e $\|\min(x_k, \nabla f(x_k) - J(x_k)^T y_k)\| \rightarrow 0$, de onde segue que

$$r(x_k, y_k) = \|(c(x_k), \min(x_k, \nabla f(x_k) - J(x_k)^T y_k))\| \rightarrow 0.$$

Assim, se $\varepsilon_k = \sqrt{nr}(x_k, y_k)$ para todo $k \in \mathcal{S}$, então

$$\|c(x_k)\| \leq r(x_k, y_k) \leq \varepsilon_k \quad \text{e} \quad (4.3)$$

$$\|\min(x_k, \nabla f(x_k) - J(x_k)^T y_k)\| \leq r(x_k, y_k) \leq \varepsilon_k. \quad (4.4)$$

Observe que (4.3) implica (4.2b). Para verificar (4.2d) podemos considerar da mesma forma que na demonstração do Lema 3.1, $x = x_k$, $u = \nabla f(x_k) - J(x_k)^T y_k$ em (3.3) e

definir w_k como em (3.3), isto é

$$[w_k]_i = \begin{cases} [\nabla f(x_k) - J(x_k)^T y_k]_i - [x_k]_i & , \text{ se } [x_k]_i \leq [\nabla f(x_k) - J(x_k)^T y_k]_i \\ 0 & , \text{ caso contrário} \end{cases}, \quad (4.5)$$

de onde segue, por (3.4), que

$$\min(x_k, \nabla f(x_k) - J(x_k)^T y_k) = \nabla f(x_k) - J(x_k)^T y_k - w_k,$$

e assim

$$\|\nabla f(x_k) - J(x_k)^T y_k - w_k\| \leq r(x_k, y_k) \leq \varepsilon_k.$$

Por construção $w_k \geq 0$, o que implica (4.2a). Para obter (4.2c), de (4.4) temos que $[x_k]_i > -r(x_k, y_k)$, logo $-[x_k]_i < r(x_k, y_k)$ e assim $\|\max(0, -x_k)\| < \sqrt{n}r(x_k, y_k)$. Finalmente se $[x_k]_i > \varepsilon_k$ então $[x_k]_i > r(x_k, y_k)$ e portanto de (4.4), $[\nabla f(x_k) - J(x_k)^T y_k]_i \leq [x_k]_i$ e assim pela definição de w_k , $[w_k]_i = 0$, o que prova (4.2e). \square

Podemos observar que o Problema MPEC (4.1) não pode ser resolvido diretamente pelo Algoritmo 5 já que este não considera restrições de desigualdade, porém usando variáveis de folga é possível transformá-lo em outro equivalente desta vez com restrições de igualdade. Primeiro, reescreveremos o Problema MPEC (4.1) como

$$\begin{aligned} \min \quad & f(x) \\ \text{s.a} \quad & g(x) \leq 0 \\ & x \geq 0 \end{aligned} \quad (4.6)$$

onde $g = (-G, -H, G^T H)$, $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e $g : \mathbb{R}^n \rightarrow \mathbb{R}^q$ o qual considerando uma variável de folga z pode-se transformar no problema

$$\begin{aligned} \min \quad & f(x) \\ \text{s.a} \quad & g(x) + z = 0 \\ & x, z \geq 0. \end{aligned} \quad (4.7)$$

Como vimos no lema anterior, o Algoritmo 5 gera pontos AKKT para o Problema (4.7), e este resultado pode ser estendido ao Problema (4.6), isto é, existem sequências

$\{x_k\} \in \mathbb{R}^n$, $\{\tilde{y}_k\} \in \mathbb{R}^p$, $\{w_k^x\} \in \mathbb{R}^n$ e $\{\tilde{\varepsilon}_k\} \in \mathbb{R}_+$ tais que $x_k \rightarrow x^*$, $\tilde{\varepsilon}_k \rightarrow 0^+$ e:

$$w_k^x \geq 0, \quad (4.8a)$$

$$\tilde{y}_k \geq 0, \quad (4.8b)$$

$$\|\max(0, g(x_k))\| \leq \tilde{\varepsilon}_k, \quad (4.8c)$$

$$\|\max(0, -x_k)\| \leq \tilde{\varepsilon}_k, \quad (4.8d)$$

$$\left\| \nabla f(x_k) + \sum_{j=1}^q [\tilde{y}_k]_j \nabla g_j(x_k) - w_k^x \right\| \leq \tilde{\varepsilon}_k, \quad (4.8e)$$

$$[w_k^x]_i = 0 \text{ se } [x_k]_i > \tilde{\varepsilon}_k, \quad (4.8f)$$

$$[\tilde{y}_k]_j = 0 \text{ se } g_j(x_k) < -\tilde{\varepsilon}_k. \quad (4.8g)$$

Proposição 4.2. *Se (x^*, z^*) é um ponto AKKT para o Problema (4.7), então x^* é um ponto AKKT para o Problema (4.6).*

Demonstração. Se (x^*, z^*) é um ponto AKKT para o Problema (4.7), então existem sequências $\{x_k\} \in \mathbb{R}^n$, $\{z_k\} \in \mathbb{R}^q$, $\{y_k\} \in \mathbb{R}^p$, $\{w_k^x\} \in \mathbb{R}^n$, $\{w_k^z\} \in \mathbb{R}^q$ e $\{\varepsilon_k\} \in \mathbb{R}_+$ tais que $x_k \rightarrow x^*$, $z_k \rightarrow z^*$, $\varepsilon_k \rightarrow 0^+$ ($\varepsilon_k = \sqrt{n}r(x_k, y_k)$) e:

$$\left\| \begin{pmatrix} \nabla f(x_k) \\ 0 \end{pmatrix} - \sum_{j=1}^q [y_k]_j \begin{pmatrix} \nabla g_j(x_k) \\ e_{n+j} \end{pmatrix} - \begin{pmatrix} w_k^x \\ w_k^z \end{pmatrix} \right\| \leq \varepsilon_k, \quad (4.9a)$$

$$\|g(x_k) + z_k\| \leq \varepsilon_k, \quad (4.9b)$$

$$\|\max(0, -z_k)\| \leq \varepsilon_k, \quad (4.9c)$$

$$\|\max(0, -x_k)\| \leq \varepsilon_k, \quad (4.9d)$$

$$[w_k^x]_i = 0 \text{ se } [x_k]_i > \varepsilon_k, \quad (4.9e)$$

$$[w_k^z]_j = 0 \text{ se } [z_k]_j > \varepsilon_k, \quad (4.9f)$$

$$w_k^x, w_k^z \geq 0, \quad (4.9g)$$

onde, por (4.5), temos

$$[w_k^x]_i = \begin{cases} [\nabla f(x_k) - \nabla g(x_k)^T y_k]_i - [x_k]_i & , \text{ se } [x_k]_i \leq [\nabla f(x_k) - \nabla g(x_k)^T y_k]_i \\ 0 & , \text{ caso contrário} \end{cases},$$

posto que neste caso $c(x, z) = g(x) + z$ e $J(x, z) = \begin{bmatrix} \nabla g(x) & I \end{bmatrix}$.

De (4.9a), temos

$$\left\| \nabla f(x_k) - \sum_{j=1}^q [y_k]_j \nabla g_j(x_k) - w_k^x \right\| \leq \varepsilon_k \quad (4.10)$$

$$\left\| [y_k]_j + [w_k^z]_j \right\| \leq \varepsilon_k \quad \forall j = 1, \dots, q. \quad (4.11)$$

Consideremos $\tilde{\varepsilon}_k = \varepsilon_k + Mq\varepsilon_k + 2\sqrt{q}\varepsilon_k$ onde $M = \sup_{j,k} \sum_{j=1}^q \|\nabla g_j(x_k)\|$. Temos que (4.9g) mostra (4.8a), (4.9d) mostra (4.8d) e para (4.8c) notemos de (4.9c) que $[z_k]_j \geq -\varepsilon_k$. Logo supondo que $g_j(x_k) > 2\varepsilon_k$, então $g_j(x_k) + [z_k]_j > \varepsilon_k$, o que contradiz (4.9b), assim $g_j(x_k) \leq 2\varepsilon_k$ e $\max(0, g_j(x_k)) \leq 2\varepsilon_k$, e assim

$$\|\max(0, g(x_k))\| \leq 2\sqrt{q}\varepsilon_k \leq \tilde{\varepsilon}_k.$$

Para (4.8b) primeiro note que de (4.11) e (4.9g) temos que

$$[y_k]_j \leq \varepsilon_k. \quad (4.12)$$

Logo,

$$\left\| \sum_{\substack{j=1 \\ |[y_k]_j| \leq \varepsilon_k}}^q [y_k]_j \nabla g_j(x_k) \right\| \leq Mq\varepsilon_k$$

de onde em (4.10)

$$\left\| \nabla f(x_k) - \sum_{\substack{j=1 \\ [y_k]_j < -\varepsilon_k}}^q [y_k]_j \nabla g_j(x_k) - w_k^x \right\| \leq \varepsilon_k + Mq\varepsilon_k \leq \tilde{\varepsilon}_k.$$

Assim, para (4.8b) definimos os multiplicadores

$$[\tilde{y}_k]_j = \begin{cases} -[y_k]_j & \text{se } [y_k]_j < -\varepsilon_k \\ 0 & \text{se } |[y_k]_j| \leq \varepsilon_k \end{cases} \quad (4.13)$$

valendo (4.8e)

$$\left\| \nabla f(x_k) + \sum_{j=1}^q [\tilde{y}_k]_j \nabla g_j(x_k) - w_k^x \right\| \leq \tilde{\varepsilon}_k.$$

Finalmente temos por (4.9e) que vale (4.8f), restando mostrar (4.8g). Para isso, consideramos $g_j(x_k) < -\tilde{\varepsilon}_k < -2\varepsilon_k$, então por (4.9b) temos que $[z_k]_j > \varepsilon_k$, o que implica de (4.9f) e (4.11) que $[w_k^z]_j = 0$ e $|[y_k]_j| \leq \varepsilon_k$, logo $[\tilde{y}_k]_j = 0$. \square

Desta proposição também podemos concluir o seguinte resultado:

Corolário 4.3. *Sob as condições da Proposição 4.2*

(1) *Se $g_j(x^*) < 0$ então para k suficientemente grande $[\tilde{y}_k]_j = 0$.*

(2) *Se $x_i^* > 0$ então $[w_k^x]_i = 0$ para todo k suficientemente grande.*

Demonstração. Para o item (1), temos que $j \notin I_g(x^*)$ logo $[z^*]_j > 0$, pois $g(x^*) + z^* = 0$. Assim como $\varepsilon_k \rightarrow 0$, temos que para k suficientemente grande $[z_k]_j > \varepsilon_k$, o que implica

por (4.9f) e (4.11) que $[w_k^z]_j = 0$ e que $|[y_k]_j| \leq \varepsilon_k$, de onde por (4.13) concluímos que para k suficientemente grande $[\tilde{y}_k]_j = 0$. O item (2) segue de (4.9e). \square

Como o Problema (4.6) é equivalente ao Problema (4.7) com variáveis de folga. A partir daqui, assumiremos que o Algoritmo 5 também resolve problemas do tipo (4.6) considerando esta equivalência.

Em [9] é demonstrado que AKKT considerando a CQ mais forte MPEC-LICQ, somente garante convergência a pontos fracamente estacionários. Porém baseados nos resultado anterior podemos fixar algumas sequências que serão de utilidade na análise de convergência em problemas MPEC. Assim resgatamos os seguintes resultados da análise feita na demonstração da proposição anterior, considerando que o conjunto $\mathcal{S} = \{k \mid (x_k, y_k) \text{ é uma V-O iteração}\}$ é infinito.

Proposição 4.4. *Seja \mathcal{S} infinito e $x^* \in \mathbb{R}^n$ um ponto de acumulação viável para o problema (4.6), da sequência $\{x_k\}_{\mathcal{S}}$ gerada pelo Algoritmo 5. Se $\varepsilon_k = \sqrt{n}r(x_k, y_k)$ e $x_k \xrightarrow{K} x^*$ para $K \subseteq \mathcal{S}$. Então $\varepsilon_k \rightarrow 0$, e*

$$\nabla f(x_k) + \sum_{j=1}^q [\hat{y}_k]_j \nabla g_j(x_k) - w_k^x \xrightarrow{K} 0 \quad (4.14)$$

onde

$$[\hat{y}_k]_j = \begin{cases} -[y_k]_j & , \text{ se } [y_k]_j < -\varepsilon_k \\ 0 & , \text{ se } |[y_k]_j| \leq \varepsilon_k \end{cases} \quad (4.15)$$

e

$$[w_k^x]_i = \begin{cases} [\nabla f(x_k) - \nabla g(x_k)^T y_k]_i - [x_k]_i & , \text{ se } [x_k]_i \leq [\nabla f(x_k) - \nabla g(x_k)^T y_k]_i \\ 0 & , \text{ caso contrário} \end{cases}, \quad (4.16)$$

Corolário 4.5. *Sob as condições da Proposição 4.4.*

- (1) $[y_k]_j \leq \varepsilon_k$ para todo $k \in K$.
- (2) Se $g_j(x^*) < 0$ então para todo $k \in K$ suficientemente grande $[\hat{y}_k]_j = 0$.
- (3) Se $x_i^* > 0$ então para todo $k \in K$ suficientemente grande $[w_k^x]_i = 0$.

A fórmula (4.15) permite identificar os multiplicadores aproximados \hat{y}_k em (4.14). Porém esta depende da sequência $\{y_k\}$ obtida do resultado da aplicação do Algoritmo 5, onde ela não aparece explicitamente, pois a atualização de y_k aparece na linha 39 do Algoritmo 5:

$$v_{k+1} = (x_{k+1}, y_{k+1}) = v_k + \alpha_k d_k + \alpha_k s_k.$$

Nesta atualização, y_k é função das direções d_k (associada a resolver um problema quadrático) e s_k (associada ao cálculo de uma direção de curvatura negativa), e ao tamanho de passo

α_k . Para resolver este problema, na próxima seção encontraremos uma fórmula explícita para os multiplicadores, fato que permitirá fechar o cálculo dos multiplicadores \hat{y}_k .

4.2 Fórmula explícita para os multiplicadores de Lagrange

No Algoritmo 5, descrito no Capítulo 2, não existe uma fórmula explícita para calcular os multiplicadores, pois os iterandos são soluções de um problema quadrático (3.34) no qual a solução é um vetor formado por variáveis primais e duais. Como nosso interesse é analisar a convergência do Algoritmo 5 quando aplicado a problemas MPEC, faremos esta análise com base no resultado da Proposição 4.4, sendo necessário mostrar como se geram os multiplicadores aproximados \hat{y}_k . Na proposição a seguir apresentamos uma fórmula explícita para a atualização dos multiplicadores de Lagrange. Para isso, consideraremos que o conjunto $\mathcal{S} = \{k \mid (x_k, y_k) \text{ é uma V-O iteração}\}$ gerado pelo Algoritmo 5 quando aplicado ao Problema (3.1) é infinito e assumiremos como válidos os resultados de convergência obtidos para o Algoritmo 5 no capítulo anterior (Lema 3.11, pag. 70).

Proposição 4.6. *Se $\{v_k\} = \{(x_k, y_k)\}$ é a sequência gerada pelo algoritmo quando é aplicado ao Problema (3.1) e $x_k \rightarrow x^*$, então para todo k suficientemente grande*

$$y_{k+1} = y_k^E - \frac{1}{\mu_k^R} c(x_k) - \frac{1}{\mu_k^R} J(x_k)(x_{k+1} - x_k).$$

Demonstração. Pelo Lema 3.11 (d), a direção de descida d_k é de descida local para todo k suficientemente grande. Isto implica que as iterações são de tipo O e V para todo k suficientemente grande e que x^* é viável. Além disso, se \hat{v}_k é a única solução do subproblema (3.34) então $d_k = \hat{v}_k - v_k$ (Algoritmo 2, linha 8), resultando do Algoritmo 5, linha 39, que

$$v_{k+1} = v_k + \alpha_k(\hat{v}_k - v_k) + \alpha_k s_k.$$

Pelo Lema 3.11 (b) temos, para todo k suficientemente grande $s_k = 0$ e $\alpha_k = 1$, resultando que

$$v_{k+1} = \hat{v}_k \quad \text{e} \quad d_k = v_{k+1} - v_k \quad \text{para todo } k \text{ suficientemente grande.}$$

Como \hat{v}_k é a única solução do Problema (3.34) então, para todo k suficientemente grande, v_{k+1} é a única solução do Problema (3.34), a saber

$$\begin{aligned} \min_v \quad & Q_k(v; y_k^E, \mu_k^R) = (v - v_k)^T \nabla M(v_k; y_k^E, \mu_k^R) + \frac{1}{2}(v - v_k)^T B(v_k; \mu_k^R)(v - v_k) \\ \text{s.a} \quad & [v]_{\mathcal{A}_\varepsilon} = 0. \end{aligned}$$

Para simplificar a análise omitiremos as variáveis das funções envolvidas, assim ∇M , B , H , J , π e ∇f serão consideradas como sendo avaliadas nas variáveis apropriadas, além

de considerar $\mu^R = \mu$. Logo aplicando as condições de otimalidade no problema anterior temos

$$\nabla M + B(v_{k+1} - v_k) - \sum_{i \in \mathcal{A}_\varepsilon} s_i e_i = 0,$$

equação que pode ser separada em termos dos conjuntos das variáveis ativas \mathcal{A}_ε e variáveis livres \mathcal{F}_ε como

$$[\nabla M + B(v_{k+1} - v_k) - s]_{\mathcal{A}_\varepsilon} = 0 \quad (4.17)$$

$$[\nabla M + B(v_{k+1} - v_k)]_{\mathcal{F}_\varepsilon} = 0. \quad (4.18)$$

Como o Lema 3.11 (a) também garante que $\mathcal{A}_\varepsilon(x_k, y_k, \mu_k^R) = \mathcal{A}(x^*)$, então $[v_{k+1}]_{\mathcal{A}_\varepsilon} = [v_k]_{\mathcal{A}_\varepsilon} = 0$ para todo k suficientemente grande, logo $[B(v_{k+1} - v_k)]_{\mathcal{F}_\varepsilon} = B_{\mathcal{F}_\varepsilon}[v_{k+1} - v_k]_{\mathcal{F}_\varepsilon}$ e assim em (4.18), temos $B_{\mathcal{F}_\varepsilon}[v_{k+1} - v_k]_{\mathcal{F}_\varepsilon} = [-\nabla M]_{\mathcal{F}_\varepsilon}$ ou, equivalentemente, considerando a definição de B (3.12), a expressão para ∇M (3.10) e a definição de v_k ,

$$\begin{pmatrix} H + \frac{2}{\mu} J^T J & J^T \\ J & \mu I \end{pmatrix}_{\mathcal{F}_\varepsilon} \begin{pmatrix} x_{k+1} - x_k \\ y_{k+1} - y_k \end{pmatrix}_{\mathcal{F}_\varepsilon} = - \begin{pmatrix} \nabla f - J^T(2\pi - y_k) \\ \mu(y_k - \pi) \end{pmatrix}_{\mathcal{F}_\varepsilon}$$

de onde

$$\begin{pmatrix} H_{\mathcal{F}_\varepsilon} + \frac{2}{\mu} J_{\mathcal{F}_\varepsilon}^T J_{\mathcal{F}_\varepsilon} & J_{\mathcal{F}_\varepsilon}^T \\ J_{\mathcal{F}_\varepsilon} & \mu I \end{pmatrix} \begin{pmatrix} [x_{k+1} - x_k]_{\mathcal{F}_\varepsilon} \\ [y_{k+1} - y_k]_{\mathcal{F}_\varepsilon} \end{pmatrix} = - \begin{pmatrix} [\nabla f - J^T(2\pi - y_k)]_{\mathcal{F}_\varepsilon} \\ \mu(y_k - \pi) \end{pmatrix}. \quad (4.19)$$

Note que

$$\begin{pmatrix} H_{\mathcal{F}_\varepsilon} + \frac{2}{\mu} J_{\mathcal{F}_\varepsilon}^T J_{\mathcal{F}_\varepsilon} & J_{\mathcal{F}_\varepsilon}^T \\ J_{\mathcal{F}_\varepsilon} & \mu I \end{pmatrix} = \begin{pmatrix} I & \frac{1}{\mu} J_{\mathcal{F}_\varepsilon}^T \\ 0 & I \end{pmatrix} \begin{pmatrix} H_{\mathcal{F}_\varepsilon} + \frac{1}{\mu} J_{\mathcal{F}_\varepsilon}^T J_{\mathcal{F}_\varepsilon} & 0 \\ 0 & \mu I \end{pmatrix} \begin{pmatrix} I & 0 \\ \frac{1}{\mu} J_{\mathcal{F}_\varepsilon} & I \end{pmatrix}.$$

Logo, usando a expressão anterior e multiplicando (4.19) à esquerda pela inversa de

$$\begin{pmatrix} I & \frac{1}{\mu} J_{\mathcal{F}_\varepsilon}^T \\ 0 & I \end{pmatrix},$$

$$\begin{aligned} & \begin{pmatrix} H_{\mathcal{F}_\varepsilon} + \frac{1}{\mu} J_{\mathcal{F}_\varepsilon}^T J_{\mathcal{F}_\varepsilon} & 0 \\ 0 & \mu I \end{pmatrix} \begin{pmatrix} I & 0 \\ \frac{1}{\mu} J_{\mathcal{F}_\varepsilon} & I \end{pmatrix} \begin{pmatrix} [x_{k+1} - x_k]_{\mathcal{F}_\varepsilon} \\ [y_{k+1} - y_k]_{\mathcal{F}_\varepsilon} \end{pmatrix} = \\ & \quad - \begin{pmatrix} I & -\frac{1}{\mu} J_{\mathcal{F}_\varepsilon}^T \\ 0 & I \end{pmatrix} \begin{pmatrix} [\nabla f - J^T(2\pi - y_k)]_{\mathcal{F}_\varepsilon} \\ \mu(y_k - \pi) \end{pmatrix} \\ \Rightarrow & \begin{pmatrix} H_{\mathcal{F}_\varepsilon} + \frac{1}{\mu} J_{\mathcal{F}_\varepsilon}^T J_{\mathcal{F}_\varepsilon} & 0 \\ 0 & \mu I \end{pmatrix} \begin{pmatrix} [x_{k+1} - x_k]_{\mathcal{F}_\varepsilon} \\ \frac{1}{\mu} J_{\mathcal{F}_\varepsilon} [x_{k+1} - x_k]_{\mathcal{F}_\varepsilon} + y_{k+1} - y_k \end{pmatrix} = \\ & \quad - \begin{pmatrix} [\nabla f - J^T(2\pi - y_k)]_{\mathcal{F}_\varepsilon} - J_{\mathcal{F}_\varepsilon}^T(y_k - \pi) \\ \mu(y_k - \pi) \end{pmatrix} \end{aligned}$$

e assim

$$\begin{pmatrix} \left(H_{\mathcal{F}_\varepsilon} + \frac{1}{\mu} J_{\mathcal{F}_\varepsilon}^T J_{\mathcal{F}_\varepsilon} \right) [x_{k+1} - x_k]_{\mathcal{F}_\varepsilon} \\ J_{\mathcal{F}_\varepsilon} [x_{k+1} - x_k]_{\mathcal{F}_\varepsilon} + \mu y_{k+1} - \mu y_k \end{pmatrix} = - \begin{pmatrix} [\nabla f]_{\mathcal{F}_\varepsilon} - J_{\mathcal{F}_\varepsilon}^T \pi \\ \mu(y_k - \pi) \end{pmatrix},$$

sistema que permite colocar em evidência y_{k+1} :

$$y_{k+1} = \pi - \frac{1}{\mu_k^R} J_{\mathcal{F}_\varepsilon} [x_{k+1} - x_k]_{\mathcal{F}_\varepsilon}.$$

Logo como $\pi(x_k, y_k^E, \mu_k^R) = y_k^E - c(x_k)/\mu_k^R$, temos para todo k suficientemente grande

$$\begin{aligned} y_{k+1} &= y_k^E - \frac{1}{\mu_k^R} c(x_k) - \frac{1}{\mu_k^R} J_{\mathcal{F}_\varepsilon} [x_{k+1} - x_k]_{\mathcal{F}_\varepsilon} \\ &= y_k^E - \frac{1}{\mu_k^R} c(x_k) - \frac{1}{\mu_k^R} J(x_k)(x_{k+1} - x_k) \end{aligned}$$

onde $J_{\mathcal{F}_\varepsilon} [x_{k+1} - x_k]_{\mathcal{F}_\varepsilon} = J(x_k)(x_{k+1} - x_k)$ pois $[x_{k+1} - x_k]_i = 0$ para índices $i \in \mathcal{A}_\varepsilon$. \square

Assumindo que o conjunto \mathcal{S} é infinito, e como válidos os resultados de convergência do Lema 3.11, obtemos da demonstração anterior uma forma de atualizar (x_k, y_k) e d_k , o qual formalizamos no seguinte corolário.

Corolário 4.7. *Se $\{v_k\} = \{(x_k, y_k)\}$ é a sequência gerada pelo algoritmo quando é aplicado ao problema Problema (3.1), \hat{v}_k é a única solução do Problema (3.34) e d_k é uma direção de descida local, então para todo k suficientemente grande*

$$v_{k+1} = \hat{v}_k \quad e \quad d_k = v_{k+1} - v_k.$$

Uma vez obtida a fórmula para a atualização dos multiplicadores do Problema (3.1) podemos obter uma fórmula equivalente para o Problema (4.7), do qual o Problema MPEC (4.1) (com variáveis de folga) é um caso particular.

Corolário 4.8. *Se (x_k, z_k, y_k) é a sequência gerada pelo Algoritmo 5 quando é aplicado ao Problema (4.7) e $(x_k, z_k) \rightarrow (x^*, z^*)$, então para todo k suficientemente grande*

$$y_{k+1} = y_k^E - \frac{1}{\mu_k^R} (g(x_k) + z_k + \nabla g(x_k)(x_{k+1} - x_k) + (z_{k+1} - z_k)). \quad (4.20)$$

Demonstração. No caso do Problema (4.7), temos que $c(x, z) = g(x) + z$ e $J(x, z) = \begin{bmatrix} \nabla g(x) & I \end{bmatrix}$, logo (4.20) segue diretamente da Proposição 4.6. \square

A fórmula (4.20) completa o resultado faltante da seção anterior e permite fechar o cálculo do \hat{y}_k em (4.15) como sendo

$$[\hat{y}_k]_j = \begin{cases} -[y_k]_j & \text{se } [y_k]_j < -\varepsilon_k \\ 0 & \text{se } |[y_k]_j| \leq \varepsilon_k \end{cases}$$

onde $\varepsilon_k = \sqrt{nr}(x_k, y_k)$ e

$$y_k = y_{k-1}^E - \frac{1}{\mu_{k-1}^R} (g(x_{k-1}) + z_{k-1} + \nabla g(x_{k-1})(x_k - x_{k-1}) + (z_k - z_{k-1}))$$

ou especificamente

$$[y_k]_i = [y_{k-1}^E]_i - \frac{1}{\mu_{k-1}^R} (g_i(x_{k-1}) + [z_{k-1}]_i + \nabla g_i(x_{k-1})^T (x_k - x_{k-1}) + [z_k - z_{k-1}]_i). \quad (4.21)$$

4.3 Propriedades de convergência do algoritmo para problemas com desigualdade

O problema MPEC (4.1) é um problema com restrições de desigualdade. Porém, considerando variáveis de folga, é um caso particular do Problema (4.7), o qual pode ser resolvido pelo Algoritmo 5. A seguir apresentamos algumas propriedades observadas quando o Algoritmo 5 é aplicado para resolver problema do tipo (4.7). Para isso veremos primeiro um resultado envolvendo o Problema (3.1). Consideraremos que o conjunto $\mathcal{S} = \{k \mid (x_k, y_k) \text{ é uma V-O iteração}\}$ gerado pelo Algoritmo 5 quando aplicado ao Problema (3.1) é infinito e assumiremos como válidos os resultados de convergência obtidos para o Algoritmo 5 no capítulo anterior (Lema 3.11, pag. 70).

Lema 4.9. *Se (x_k, y_k) é a sequência gerada pelo Algoritmo 5 quando é aplicado ao Problema (3.1) e $x_k \rightarrow x^*$, então*

$$\frac{1}{\mu_k^R} \|x_{k+1} - x_k\| \rightarrow 0 \quad e \quad (4.22)$$

$$\frac{1}{\mu_k^R} \|x_k - x^*\| \rightarrow 0. \quad (4.23)$$

Demonstração. Temos pelo Lema 3.11 (d) que para todo k suficientemente grande a direção de descida local $d_k = (p_k, q_k)$ é computada por (3.34) e satisfaz $\max\{\|p_k\|, \|q_k\|\} \leq O(r(x_k, y_k))$. Logo para todo k suficientemente grande existe uma constante $L > 0$ tal que

$$\|p_k\| \leq Lr(x_k, y_k).$$

Pelo Corolário 4.7 temos que para todo k suficientemente grande, $d_k = v_{k+1} - v_k$ (com $v_k = (x_k, y_k)$), logo

$$\|x_{k+1} - x_k\| \leq Lr(x_k, y_k).$$

Agora pelo Lema 3.11 (c) temos que para todo k suficientemente grande $r(x_k, y_k)^{\bar{\gamma}} = \mu_k^R$ com $\bar{\gamma} \in (0, 1)$, logo

$$\|x_{k+1} - x_k\| \leq L (\mu_k^R)^{1/\bar{\gamma}}$$

e assim

$$\frac{\|x_{k+1} - x_k\|}{\mu_k^R} \leq L (\mu_k^R)^{\frac{1}{\bar{\gamma}} - 1}.$$

Pelo Teorema 3.8, $\mu_k^R \rightarrow 0$, e como $\frac{1}{\bar{\gamma}} - 1 > 0$ então

$$\frac{\|x_{k+1} - x_k\|}{\mu_k^R} \rightarrow 0. \quad (4.24)$$

Por outro lado temos pelo Corolário 4.7 que para todo k suficientemente grande $v_{k+1} = \hat{v}_k$. Logo pelo Lema 3.11 (e) temos que para todo k suficientemente grande existe $N > 0$ tal que

$$\|x_{k+1} - x^*\| = \|\hat{x}_k - x^*\| \leq Nr(x_k, y_k)^{1+\bar{\gamma}}.$$

Logo pelo Lema 3.11 (c) temos que para todo k suficientemente grande

$$\|x_{k+1} - x^*\| \leq N (\mu_k^R)^{\frac{1+\bar{\gamma}}{\bar{\gamma}}} = N (\mu_k^R)^{1+\frac{1}{\bar{\gamma}}}$$

e como $\mu_k^R \rightarrow 0$ pelo Teorema 3.8,

$$\frac{\|x_{k+1} - x^*\|}{\mu_k^R} \leq N (\mu_k^R)^{\frac{1}{\bar{\gamma}}} \rightarrow 0. \quad (4.25)$$

Finalmente de (4.24) e (4.25) concluímos que

$$\frac{\|x_k - x^*\|}{\mu_k^R} \leq \frac{\|x_k - x_{k+1}\|}{\mu_k^R} + \frac{\|x_{k+1} - x^*\|}{\mu_k^R} \rightarrow 0.$$

□

Aplicando o Lema anterior a problemas do tipo (4.7), obtemos o seguinte resultado:

Corolário 4.10. *Seja (x_k, z_k, y_k) a sequência gerada pelo Algoritmo 5 quando aplicado ao problema Problema (4.7) e $(x_k, z_k) \rightarrow (x^*, z^*)$, então*

$$\frac{1}{\mu_k^R} (x_{k+1} - x_k) \rightarrow 0, \quad (4.26)$$

$$\frac{1}{\mu_k^R} (z_{k+1} - z_k) \rightarrow 0, \quad (4.27)$$

$$\frac{[z_k]_i}{\mu_k^R} \rightarrow 0 \text{ se } g_i(x^*) = 0, \quad (4.28)$$

$$\frac{1}{\mu_k^R} (g_i(x_{k+1}) - g_i(x_k)) \rightarrow 0. \quad (4.29)$$

Demonstração. Temos pelo Lema 4.9, considerando (x_k, z_k) no lugar de x_k , que (4.26) e (4.27) seguem diretamente de (4.22). Agora como (x^*, z^*) é um ponto viável do Problema (4.7), se $g_i(x^*) = 0$, então $[z^*]_i = 0$, logo por (4.23) concluímos (4.28). Finalmente como

g_i é contínua, segue do Teorema de Valor Médio que existe ξ_k pertencente ao segmento de reta $[x_{k-1}, x_k]$ tal que

$$\frac{1}{\mu_k^R} (g_i(x_{k+1}) - g_i(x_k)) = \frac{1}{\mu_k^R} (\nabla g_i(\xi_k)^T (x_{k+1} - x_k))$$

e como $\xi_k \rightarrow x^*$, temos por (4.26) que vale (4.29). \square

4.4 Resultados de convergência para MPECs

Uma vez identificada uma forma explícita (4.21) que permite a análise dos multiplicadores \hat{y}_k (4.15) para o Problema (4.6) (Proposição 4.4), podemos aplicar este resultado ao Problema MPEC (4.1). Para poder utilizar os resultados das seções anteriores, consideraremos $\mathcal{S} = \{k \mid (x_k, y_k) \text{ é uma V-O iteração}\}$ infinito, $\varepsilon_k = \sqrt{nr}(x_k, y_k)$ ($\varepsilon_k \rightarrow 0$) e x^* um ponto de acumulação viável para o problema (4.6) da sequência $\{x_k\}_{\mathcal{S}}$ gerada pelo Algoritmo 5, isto é, $x_k \xrightarrow{K} x^*$ para $K \subseteq \mathcal{S}$. Além disso consideraremos como válidos os resultados de convergência obtidos para o Algoritmo 5 no capítulo anterior (Lema 3.11, pag. 70).

Para o caso do Problema MPEC (4.1), temos em (4.21) que

$$g(x) = (-G(x), -H(x), G(x)^T H(x)),$$

e para uma melhor análise consideraremos, em (4.15) e (4.21), $\hat{y}_k = (\hat{y}_k^G, \hat{y}_k^H, \hat{y}_k^0)$, $z_k = (z_k^G, z_k^H, z_k^0)$ e $y_k^E = (y_k^{E,G}, y_k^{E,H}, y_k^{E,0})$ para destacar a relação que possuem as variáveis com as restrições. Assim, se para k suficientemente grande

$$[y_{k-1}^{E,G}]_i + \frac{1}{\mu_{k-1}^R} G_i(x_{k-1}) - \underbrace{\frac{1}{\mu_{k-1}^R} ([z_{k-1}^G]_i - \nabla G_i(x_{k-1})^T (x_k - x_{k-1}) + [z_k^G - z_{k-1}^G]_i)}_{F_k^{G_i}} < -\varepsilon_k \quad (4.30)$$

então

$$[\hat{y}_k^G]_i = -[y_{k-1}^{E,G}]_i - \frac{1}{\mu_{k-1}^R} G_i(x_{k-1}) + F_k^{G_i}, \quad (4.31)$$

e se

$$[y_{k-1}^{E,H}]_i + \frac{1}{\mu_{k-1}^R} H_i(x_{k-1}) - \underbrace{\frac{1}{\mu_{k-1}^R} ([z_{k-1}^H]_i - \nabla H_i(x_{k-1})^T (x_k - x_{k-1}) + [z_k^H - z_{k-1}^H]_i)}_{F_k^{H_i}} < -\varepsilon_k \quad (4.32)$$

então

$$[\hat{y}_k^H]_i = -[y_{k-1}^{E,H}]_i - \frac{1}{\mu_{k-1}^R} H_i(x_{k-1}) + F_k^{H_i}. \quad (4.33)$$

Note que pelo Corolário 4.10

$$F_k^{G_i} \rightarrow 0 \quad \text{e} \quad F_k^{H_i} \rightarrow 0. \quad (4.34)$$

Temos pela Proposição 4.4, quando aplicado ao Problema MPEC (4.1), que se $x_k \xrightarrow{K} x^*$ e $\varepsilon_k = \sqrt{n}r(x_k, y_k)$ então $\varepsilon_k \rightarrow 0$ e

$$\nabla f(x_k) - \nabla G(x_k)^T \hat{y}_k^G - \nabla H(x_k)^T \hat{y}_k^H + \hat{y}_k^0 (\nabla G(x_k)^T H(x_k) + \nabla H(x_k)^T G(x_k)) - w_k^x \xrightarrow{K} 0 \quad (4.35)$$

onde $\hat{y}_k = (y_k^G, y_k^H, y_k^0)$ é definido por (4.15) e (4.21), e w_k^x por (4.16). Especificamente neste caso, recuperando (4.16), temos

$$[w_k^x]_i = \begin{cases} [\nabla f(x_k) - \nabla g(x_k)^T y_k]_i - [x_k]_i & , \text{ se } [x_k]_i \leq [\nabla f(x_k) - \nabla g(x_k)^T y_k]_i \\ 0 & , \text{ caso contrário} \end{cases} \quad (4.36)$$

onde

$$\nabla g(x) = \begin{bmatrix} -\nabla G(x) \\ -\nabla H(x) \\ (\nabla G^T(x)H(x) + \nabla H^T(x)G(x))^T \end{bmatrix}.$$

Em particular, caso a sequência $\{\hat{y}_k\}_K$ seja limitada temos $\hat{y}_k \xrightarrow{K_1} y^*$ ($S \subseteq K$) para alguma subsequência. Nesse caso, por (4.36), temos $w_k^x \xrightarrow{K_1} w^*$ com

$$[w^*]_i = \begin{cases} [\nabla f(x^*) - \nabla g(x^*)^T y^*]_i - [x^*]_i & , \text{ se } [x^*]_i \leq [\nabla f(x^*) - \nabla g(x^*)^T y^*]_i \\ 0 & , \text{ caso contrário.} \end{cases}$$

Assim x^* é KKT com os multiplicadores y^* e w^* , e pela Proposição 2.19, x^* é S -estacionário.

Em geral, notemos primeiro que se $[x^*]_i > 0$, temos pelo Corolário 4.5 (3) que para todo k suficientemente grande $[w_k^x]_i = 0$. Logo (4.35) pode ser escrita como

$$\nabla f(x_k) - \nabla G(x_k)^T (\hat{y}_k^G - \hat{y}_k^0 H(x_k)) - \nabla H(x_k)^T (\hat{y}_k^H - \hat{y}_k^0 G(x_k)) - \sum_{i \in \mathcal{A}(x^*)} [w_k^x]_i e_i = \tilde{\varepsilon}_k$$

com $\tilde{\varepsilon}_k \xrightarrow{K} 0$, e se consideramos

$$\begin{aligned} \gamma^k &= \hat{y}_k^G - \hat{y}_k^0 H(x_k) \quad \text{e} \\ \nu^k &= \hat{y}_k^H - \hat{y}_k^0 G(x_k), \end{aligned} \quad (4.37)$$

temos

$$\nabla f(x_k) - \nabla G(x_k)^T \gamma^k - \nabla H(x_k)^T \nu^k - \sum_{i \in \mathcal{A}(x^*)} [w_k^x]_i e_i = \tilde{\varepsilon}_k. \quad (4.38)$$

4.4.1 Convergência para pontos S-estacionários

Apresentamos um resultado de convergência para pontos S-estacionários utilizando uma abordagem similar à demonstração feita em [1] para métodos baseados em Lagrangiano Aumentado. Para isso consideraremos a condição MPEC-RCPLD da Definição 2.31, que no caso do Problema (4.1) fica da seguinte forma:

Se $\mathcal{I}_G \subset I_{0+}$, $\mathcal{I}_H \subset I_{+0}$ são tais que

$$\{\nabla G_i(x^*), \nabla H_j(x^*) \mid i \in \mathcal{I}_G, j \in \mathcal{I}_H\}$$

é uma base para o conjunto

$$\text{span}\{\nabla G_i(x^*), \nabla H_j(x^*) \mid i \in I_{0+}, j \in I_{+0}\},$$

x^* satisfaz MPEC-RCPLD se existe uma vizinhança $B(x^*)$ de x^* tal que:

- $\{\nabla G_i(x), \nabla H_j(x) \mid i \in I_{0+}, j \in I_{+0}\}$ possui o mesmo posto para todo $x \in B(x^*)$,
- para cada $\tilde{I} \subset \mathcal{A}(x^*)$ e $\tilde{I}_G, \tilde{I}_H \subset I_{00}$, quando existem multiplicadores w^x , λ^G , λ^H , não todos nulos, satisfazendo $(\lambda_i^G \lambda_i^H = 0 \text{ ou } \lambda_i^G, \lambda_i^H > 0) \forall i \in I_{00}$ e

$$\sum_{i \in \mathcal{I}_G \cup \tilde{I}_G} \lambda_i^G \nabla G_i(x^*) + \sum_{j \in \mathcal{I}_H \cup \tilde{I}_H} \lambda_j^H \nabla H_j(x^*) + \sum_{k \in \tilde{I}} w_k^x e_k = 0$$

então

$$\{\nabla G_i(x), \nabla H_j(x), e_k \mid i \in \mathcal{I}_G \cup \tilde{I}_G, j \in \mathcal{I}_H \cup \tilde{I}_H, k \in \tilde{I}\}$$

é linearmente dependente para cada $x \in B(x^*)$.

Destacamos que para usar a condição MPEC-RCPLD na análise da convergência será preciso utilizar dois resultados prévios. O primeiro é devido a Izmailov, Solodov e Uskov, e para uma melhor compreensão detalhamos a demonstração.

Lema 4.11. [40, Lema 3.1] *Sejam $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e $h : \mathbb{R}^n \rightarrow \mathbb{R}^l$ diferenciáveis em uma vizinhança de $x^* \in \mathbb{R}^n$ com derivadas contínuas em x^* . Se $\{\nabla h_i(x^*) \mid i = 1, \dots, l\}$ é linearmente independente, $x_k \rightarrow x^*$, $\alpha_k \rightarrow \alpha^*$, $\beta_k \rightarrow 0$ e*

$$\alpha_k \nabla f(x_k) - \nabla h(x_k)^T \lambda_k = \beta_k \tag{4.39}$$

então existe um único λ^ tal que $\lambda_k \rightarrow \lambda^*$ e*

$$\alpha^* \nabla f(x^*) - \nabla h(x^*)^T \lambda^* = 0. \tag{4.40}$$

Em particular, se $\alpha^ = 0$, então $\lambda^* = 0$.*

Demonstração. Suponhamos que a sequência $\{\lambda_k\}$ é ilimitada. Então existe um conjunto infinito $K \in \mathbb{N}$ tal que $\|\lambda_k\| \xrightarrow{K} +\infty$. Considerando agora um subsequência convergente $\left\{\frac{\lambda_k}{\|\lambda_k\|}\right\}_{k \in K_1}$ com $K_1 \subseteq K$ e $\frac{\lambda_k}{\|\lambda_k\|} \xrightarrow{K_1} \mu$ temos que $\|\mu\| = 1$. Considerando $k \in K_1$ em (4.39) temos que

$$\frac{\alpha_k}{\|\lambda_k\|} \nabla f(x_k) - \nabla h(x^k)^T \frac{\lambda_k}{\|\lambda_k\|} = \frac{\beta_k}{\|\lambda_k\|} \text{ para } k \in K_1$$

e passando ao limite, $\nabla h(x^*)^T \mu = 0$, o que contradiz o fato dos gradientes serem linearmente independentes e μ ter norma 1. Assim $\{\lambda_k\}$ é limitada, logo considerando qualquer subsequência convergente de $\{\lambda_k\}$ obtemos (4.40) para algum $\lambda^* \in \mathbb{R}^l$. Para demonstrar a unicidade, considere $\bar{\lambda} \in \mathbb{R}^l$ limite de outra subsequência de $\{\lambda_k\}$ tal que

$$\alpha^* \nabla f(x^*) - \nabla h(x^*)^T \bar{\lambda} = 0,$$

logo, comparando com (4.40), $\nabla h(x^*)^T \bar{\lambda} = \nabla h(x^*)^T \lambda^*$. Como $\{\nabla h_i(x^*) \mid i = 1, \dots, l\}$ é linearmente independente então $\lambda^* = \bar{\lambda}$. Assim $\lambda_k \rightarrow \lambda^*$. \square

O segundo se baseia em um resultado de Andreani, Haeser, Schuverdt e Silva [8, Lema 1].

Lema 4.12. *Seja $x = \sum_{i=1}^{m+p} \alpha_i v_i$ com $v_i \in \mathbb{R}^n$ e $\alpha_i \in \mathbb{R}$. Se $\{v_i\}_{i \in \{1, \dots, m\}}$ é linearmente independente e $\alpha_i \leq 0$ para $i > m$, então existe $J \subset \{m+1, \dots, m+p\}$ e escalares $\tilde{\alpha}_i$ tais que*

- $x = \sum_{i \in \{1, \dots, m\} \cup J} \tilde{\alpha}_i v_i$,
- $\tilde{\alpha}_i \leq 0$ para todo $i \in J$,
- $\{v_i\}_{i \in \{1, \dots, m\} \cup J}$ é linearmente independente.

Demonstração. Sem perda de generalidade suporemos que $\{v_i\}_{i \in \{1, \dots, m+p\}}$ é linearmente dependente, pois caso contrário o resultado seria imediato. Logo, existe $\beta \in \mathbb{R}^{m+p}$ não nulo tal que

$$\sum_{i=1}^{m+p} \beta_i v_i = 0. \quad (4.41)$$

Assim podemos escrever x como

$$x = \sum_{i=1}^{m+p} (\alpha_i - \gamma \beta_i) v_i, \text{ para todo } \gamma \in \mathbb{R}. \quad (4.42)$$

Como $\beta \neq 0$, sem perda de generalidade por (4.41) podemos considerar que existe $\beta_i > 0$ para algum $i > m$. Assim se

$$\bar{\gamma} = \max\left\{\frac{\alpha_i}{\beta_i} \mid \beta_i > 0, i > m\right\},$$

$L = \{i \mid \beta_i > 0, i > m \text{ e } \alpha_i - \bar{\gamma}\beta_i = 0\}$ e $J = \{m+1, \dots, m+p\} \setminus L$ temos de (4.42)

$$\begin{aligned} x &= \sum_{i \in \{1, \dots, m+p\} \setminus L} (\alpha_i - \bar{\gamma}\beta_i)v_i, \\ x &= \sum_{i \in \{1, \dots, m\} \cup J} (\alpha_i - \bar{\gamma}\beta_i)v_i. \end{aligned} \quad (4.43)$$

Podemos expressar assim x como uma combinação linear dos v_i com pelo menos um vetor v_j a menos, para $j > m$. Afirmamos que se $i > m$ então $\alpha_i - \bar{\gamma}\beta_i \leq 0$. De fato se $\beta_i > 0$ temos que $\bar{\gamma} \geq \frac{\alpha_i}{\beta_i}$ de onde $\alpha_i - \bar{\gamma}\beta_i \leq 0$. Caso contrário, se $\beta_i \leq 0$ então $\bar{\gamma}\beta_i \geq 0$, pois $\bar{\gamma} \leq 0$ por construção, assim $\alpha_i - \bar{\gamma}\beta_i \leq 0$. Finalmente podemos repetir estes argumentos até conseguir que os v_i em (4.43) sejam linearmente independentes. \square

A seguir apresentamos um dos resultados principais desta tese, no qual mostramos convergência a pontos S-estacionários. Este resultado é análogo ao obtido em [1] para métodos baseados em Lagrangiano Aumentado e da mesma forma que em [1], exigiremos válida a condição MPEC-RCPLD e que a sequência de multiplicadores aproximados associada à restrição de complementariedade seja limitada.

Teorema 4.13. *Seja x^* um ponto de acumulação viável para o Problema (4.1) da sequência $\{x_k\}_S$ gerada pelo Algoritmo 5. Se $\liminf \frac{1}{\mu_{k-1}^R} G(x_{k-1})^T H(x_{k-1}) < \infty$ e x^* satisfaz a condição MPEC-RCPLD então x^* é S-estacionário.*

Demonstração. Temos, de (4.38), que se $x_k \xrightarrow{K} x^*$ para $K \subseteq S$, então

$$\nabla f(x_k) - \nabla G(x_k)^T \gamma^k - \nabla H(x_k)^T \nu^k - \sum_{i \in \mathcal{A}(x^*)} [w_k^x]_i e_i \xrightarrow{K} 0 \quad (4.44)$$

onde por (4.36) e (4.37), $[w_k^x]_i \geq 0$ e

$$\begin{aligned} \gamma^k &= \hat{y}_k^G - \hat{y}_k^0 H(x_k), \\ \nu^k &= \hat{y}_k^H - \hat{y}_k^0 G(x_k). \end{aligned} \quad (4.45)$$

Como $\liminf \frac{1}{\mu_{k-1}^R} G(x_{k-1})^T H(x_{k-1}) < \infty$ então existe um conjunto infinito $K_1 \subset K$ onde $\frac{1}{\mu_{k-1}^R} G(x_{k-1})^T H(x_{k-1})$ é limitado. Temos de (4.15) e (4.21) que para todo k suficientemente grande $\hat{y}_k^0 = 0$ ou

$$\begin{aligned} \hat{y}_k^0 &= -y_{k-1}^{E,0} + \frac{1}{\mu_{k-1}^R} (G(x_{k-1})^T H(x_{k-1}) + z_{k-1}^0 + (\nabla G^T(x_{k-1}) H(x_{k-1}) \\ &\quad + \nabla H^T(x_{k-1}) G(x_{k-1}))^T (x_k - x_{k-1}) + (z_k^0 - z_{k-1}^0)). \end{aligned}$$

Como neste caso todos os iterandos são de tipo O e V, temos pela condição (3.23) que $\{y_{k-1}^{E,0}\}$ é limitada. Logo pelo Corolário 4.10, $\{\hat{y}_k^0\}$ é limitado em K_1 . Por outro lado se $H_i(x^*) > 0$, então pelo Corolário 4.5 (2), $\hat{y}_k^H = 0$ para todo $k \in K$ suficientemente grande,

e similarmente se $G_i(x^*) > 0$ então $\hat{y}_k^G = 0$ para todo $k \in K$ suficientemente grande. Assim em (4.45)

$$\begin{aligned} \gamma_{I_{+0}}^k &\xrightarrow{K_1} 0 \quad \text{e} \\ \nu_{I_{0+}}^k &\xrightarrow{K_1} 0. \end{aligned}$$

Logo (4.44) pode ser escrito como

$$\nabla f(x_k) - \sum_{i \in I_{00} \cup I_{0+}} \gamma_i^k \nabla G_i(x_k) - \sum_{i \in I_{00} \cup I_{+0}} \nu_i^k \nabla H_i(x_k) - \sum_{i \in \mathcal{A}(x^*)} [w_k^x]_i e_i \xrightarrow{K_1} 0. \quad (4.46)$$

Afirmamos que em (4.46) podemos considerar que

$$\text{se } i \in I_{00}, \text{ então } \gamma_i^k \geq 0 \text{ e } \nu_i^k \geq 0 \text{ para } k \in K_1 \text{ suficientemente grande,} \quad (4.47)$$

caso contrário, como $[\hat{y}_k^G]_i \geq 0$ e $[\hat{y}_k^H]_i \geq 0$, teríamos de (4.45) que existiria um conjunto infinito $K_2 \subset K_1$, tal que $\gamma_i^k < 0$ ou $\nu_i^k < 0$ onde

$$\begin{aligned} 0 &\leq [\hat{y}_k^G]_i < \hat{y}_k^0 H(x_k) \xrightarrow{K_2} 0 \quad \text{ou} \\ 0 &\leq [\hat{y}_k^H]_i < \hat{y}_k^0 G(x_k) \xrightarrow{K_2} 0. \end{aligned}$$

Assim, por (4.45),

$$\begin{aligned} \gamma_i &\xrightarrow{K_2} 0 \quad \text{ou} \\ \nu_i &\xrightarrow{K_2} 0, \end{aligned}$$

o que permitiria redefinir γ_i^k ou ν_i^k em K_2 como sendo seus opostos positivos e satisfazendo

$$\nabla f(x_k) - \sum_{i \in I_{00} \cup I_{0+}} \gamma_i^k \nabla G_i(x_k) - \sum_{i \in I_{00} \cup I_{+0}} \nu_i^k \nabla H_i(x_k) - \sum_{i \in \mathcal{A}(x^*)} [w_k^x]_i e_i \xrightarrow{K_2} 0.$$

Consideremos agora $\mathcal{I}_G \subset I_{0+}$, $\mathcal{I}_H \subset I_{+0}$ como na definição da condição de qualificação MPEC-RCPLD, tais que

$$\mathcal{D}(x^*) = \{\nabla G_i(x^*), \nabla H_j(x^*) \mid i \in \mathcal{I}_G, j \in \mathcal{I}_H\}$$

é uma base para o conjunto

$$\mathcal{E}(x^*) = \text{span}\{\nabla G_i(x^*), \nabla H_j(x^*) \mid i \in I_{0+}, j \in I_{+0}\}.$$

Da primeira condição de MPEC-RCPLD, temos que $\mathcal{D}(x_k)$ é uma base para $\mathcal{E}(x_k)$ para todo $k \in K_1$ suficientemente grande, digamos $k \in K_3 \subset K_1$. Logo existem sequências

$\{\hat{\gamma}_{\mathcal{I}_G}^k\}$ e $\{\hat{\nu}_{\mathcal{I}_H}^k\}$ com $k \in K_3$ tais que

$$\sum_{i \in I_{0+}} \gamma_i^k \nabla G_i(x_k) + \sum_{i \in I_{+0}} \nu_i^k \nabla H_i(x_k) = \sum_{i \in \mathcal{I}_G} \hat{\gamma}_i^k \nabla G_i(x_k) + \sum_{i \in \mathcal{I}_H} \hat{\nu}_i^k \nabla H_i(x_k)$$

e assim, de (4.46),

$$\begin{aligned} & \nabla f(x_k) - \sum_{i \in \mathcal{I}_G} \hat{\gamma}_i^k \nabla G_i(x_k) - \sum_{i \in \mathcal{I}_H} \hat{\nu}_i^k \nabla H_i(x_k) - \\ & \sum_{i \in I_{00}} \gamma_i^k \nabla G_i(x_k) - \sum_{i \in I_{00}} \nu_i^k \nabla H_i(x_k) - \sum_{i \in \mathcal{A}(x^*)} [w_k^x]_i e_i \xrightarrow{K_3} 0. \end{aligned}$$

Os gradientes dos dois primeiros somatórios são linearmente independentes e os multiplicadores das três ultimas são não negativos. Logo considerando o Lema 4.12 na última relação, temos que para cada $k \in K_3$ existem conjuntos $I_0^k \subset I_{00}$ e $A_0^k \subset \mathcal{A}(x^*)$, e multiplicadores $\hat{\gamma}_{I_0^k}^k, \hat{\nu}_{I_0^k}^k, [\hat{w}_k^x]_{A_0} \geq 0$ tais que

$$\begin{aligned} & \nabla f(x_k) - \sum_{i \in \mathcal{I}_G} \hat{\gamma}_i^k \nabla G_i(x_k) - \sum_{i \in \mathcal{I}_H} \hat{\nu}_i^k \nabla H_i(x_k) - \\ & \sum_{i \in I_0^k} \hat{\gamma}_i^k \nabla G_i(x_k) - \sum_{i \in I_0^k} \hat{\nu}_i^k \nabla H_i(x_k) - \sum_{i \in A_0^k} [\hat{w}_k^x]_i e_i \xrightarrow{K_3} 0, \end{aligned}$$

onde

$$\{\nabla G_i(x_k), \nabla H_j(x_k), e_l \mid i \in \mathcal{I}_G \cup I_0^k, j \in \mathcal{I}_H \cup I_0^k, l \in A_0^k\}$$

é um conjunto linearmente independente e $\hat{\gamma}_{I_0^k}^k, \hat{\nu}_{I_0^k}^k, [\hat{w}_k^x]_{A_0^k} \geq 0$. Como somente existe um número finito de conjuntos I_0^k e A_0^k , então existe um conjunto infinito $K_4 \subset K_3$ e conjuntos $\hat{I}_0 \subset I_{00}$ e $\hat{A}_0 \subset \mathcal{A}(x^*)$ tais que

$$\nabla f(x_k) - \sum_{i \in \mathcal{I}_G \cup \hat{I}_0} \hat{\gamma}_i^k \nabla G_i(x_k) - \sum_{i \in \mathcal{I}_H \cup \hat{I}_0} \hat{\nu}_i^k \nabla H_i(x_k) - \sum_{i \in \hat{A}_0} [\hat{w}_k^x]_i e_i \xrightarrow{K_4} 0$$

e o conjunto

$$\{\nabla G_i(x_k), \nabla H_j(x_k), e_l \mid i \in \mathcal{I}_G \cup \hat{I}_0, j \in \mathcal{I}_H \cup \hat{I}_0, l \in \hat{A}_0\}$$

é linearmente independente. Pelo Lema 4.11, existem $\gamma_{\mathcal{I}_G \cup \hat{I}_0}^*, \nu_{\mathcal{I}_H \cup \hat{I}_0}^*$ e $w_{\hat{A}_0}^*$ tais que

$$\hat{\gamma}_{\mathcal{I}_G \cup \hat{I}_0}^k \xrightarrow{K_4} \gamma_{\mathcal{I}_G \cup \hat{I}_0}^*, \hat{\nu}_{\mathcal{I}_H \cup \hat{I}_0}^k \xrightarrow{K_4} \nu_{\mathcal{I}_H \cup \hat{I}_0}^* \quad \text{e} \quad [\hat{w}_k^x]_{\hat{A}_0} \xrightarrow{K_4} w_{\hat{A}_0}^*$$

e $\gamma_{\hat{I}_0}^*, \nu_{\hat{I}_0}^*, w_{\hat{A}_0}^* \geq 0$. Finalmente considerando os multiplicadores restantes iguais a zero concluímos que x^* é S -estacionário. \square

O resultado anterior pede que $\liminf \frac{1}{\mu_{k-1}^R} G(x_{k-1})^T H(x_{k-1}) < \infty$. A seguir deixaremos essa exigência de lado. Por causa disso, consideraremos uma condição de

qualificação mais forte (MPEC-LICQ), porém ganharemos convergência a pontos M-estacionários (mais fraco que S-estacionário).

4.4.2 Convergência para pontos M-estacionários

Consideraremos agora que MPEC-LICQ é satisfeita em x^* e mostraremos convergência para pontos M-estacionários. O resultado está baseado na demonstração feita em [40], para métodos baseados em Lagrangiano Aumentado. Primeiramente será assumido sem perda de generalidade que $G(x^*) = 0$, e para a análise consideraremos os seguintes conjuntos:

$$\begin{aligned} I_0 &= I_{00} = \{i : H_i(x^*) = 0\} \text{ e} \\ I_+ &= I_{0+} = \{i : H_i(x^*) > 0\}. \end{aligned}$$

Assim, (4.38) pode ser reescrita como

$$\nabla f(x_k) - \nabla G(x_k)^T \gamma^k - \nabla H_{I_0}(x_k)^T \nu_{I_0}^k - \nabla H_{I_+}(x_k)^T \nu_{I_+}^k - \sum_{i \in \mathcal{A}(x^*)} [w_k^x]_i e_i = \tilde{\varepsilon}_k \quad (4.48)$$

e a condição MPEC-LICQ em x^* para o Problema (4.1) neste caso significa que o conjunto $\{\nabla G_i(x_k) : i = 1, \dots, n\} \cup \{\nabla H_i(x_k) : i \in I_0\} \cup \{e_i : i \in \mathcal{A}(x^*)\}$ é linearmente independente.

Para ver que o ponto x^* é M-estacionário consideraremos ao igual que na seção anterior que $x_k \xrightarrow{K} x^*$ para $K \subseteq \mathcal{S}$. Suponhamos primeiro que $\{\hat{y}_k^0 \mid k \in K_1\}$ é limitada para algum subconjunto infinito $K_1 \subseteq K$. Lembrando que se $i \in I_+$, então pelo Corolário 4.5 (2), $[\hat{y}_k^H]_i = 0$ para $k \in K_1$ suficientemente grande. Assim pela definição de ν^k (4.37) concluímos que

$$\nabla H_{I_+}(x_k)^T \nu_{I_+}^k = -\hat{y}_k^0 \nabla H_{I_+}(x_k)^T G_{I_+}(x_k) \xrightarrow{K_1} 0.$$

Agora, se vale MPEC-LICQ em x^* , então o conjunto $\{\nabla G_i(x_k) : i = 1, \dots, n\} \cup \{\nabla H_i(x_k) : i \in I_0\} \cup \{e_i : i \in \mathcal{A}(x^*)\}$ é linearmente independente. Logo aplicando o Lema 4.11 em (4.48) com $\beta_k = \nabla H_{I_+}(x_k)^T \nu_{I_+}^k + \varepsilon_k$ e $\alpha_k = 1$ temos que existem γ^* , $\nu_{I_0}^*$ e $[\bar{w}^*]_i$ para $i \in \mathcal{A}(x^*)$ tais que

$$\nabla f(x^*) - \nabla G(x^*)^T \gamma^* - \nabla H_{I_0}(x^*)^T \nu_{I_0}^* - \sum_{i \in \mathcal{A}(x^*)} [\bar{w}^*]_i e_i = 0$$

e se definimos

$$w^* = \begin{cases} [\bar{w}^*]_i & \text{se } i \in \mathcal{A}(x^*) \\ 0 & \text{se } i \in \mathcal{F}(x^*) \end{cases}$$

então

$$\nabla f(x^*) - \nabla G(x^*)^T \gamma^* - \nabla H_{I_0}(x^*)^T \nu_{I_0}^* - w^* = 0$$

mostrando assim que x^* é S -estacionário.

Para o caso em que $\hat{y}_k^0 \xrightarrow{K} +\infty$ veremos alguns resultados prévios, relacionados às sequências obtidas pelo Algoritmo 5, as quais apresentamos no lema a seguir. Durante todo o resto desta seção consideraremos que $\hat{y}_k^0 \xrightarrow{K} +\infty$.

Lema 4.14. *Seja x^* um ponto de acumulação viável para o Problema (4.1) da sequência $\{x_k\}_S$ gerada pelo Algoritmo 5. Se x^* satisfaz as condições de MPEC-LICQ, $x_k \xrightarrow{K} x^*$ para $K \subseteq \mathcal{S}$ e $\hat{y}_k^0 \xrightarrow{K} +\infty$, então*

- (a) $\hat{y}_k^0 \mu_{k-1}^R \xrightarrow{K} 0$;
- (b) Se $I_+ \neq \emptyset$, então $\frac{\gamma_k^k}{\hat{y}_k^0} \xrightarrow{K} 0$ e $\frac{\nu_{I_0}^k}{\hat{y}_k^0} \xrightarrow{K} 0$;
- (c) $I_0 \neq \emptyset$;
- (d) $\frac{1}{\hat{y}_k^0 \mu_{k-1}^R} G_{I_0}(x_{k-1})^T H_{I_0}(x_{k-1}) \xrightarrow{K} M > 0$;
- (e) $\frac{\nu_{I_+}^k}{(\hat{y}_k^0)^2 \mu_{k-1}^R} \xrightarrow{K} H_{I_+}(x^*)$.

Demonstração. Como neste caso

$$J(x, z) = \begin{bmatrix} -\nabla G(x) & I & 0 & 0 \\ -\nabla H(x) & 0 & I & 0 \\ (\nabla G^T(x)H(x) + \nabla H^T(x)G(x))^T & 0 & 0 & 1 \end{bmatrix}$$

temos por (4.15) e (4.21) que para todo $k \in K$ suficientemente grande

$$\begin{aligned} \hat{y}_k^0 &= -y_{k-1}^{E,0} + \frac{1}{\mu_{k-1}^R} (G(x_{k-1})^T H(x_{k-1}) + z_{k-1}^0 + (\nabla G(x_{k-1})^T H(x_{k-1}) \\ &\quad + \nabla H(x_{k-1})^T G(x_{k-1}))^T (x_k - x_{k-1}) + (z_k^0 - z_{k-1}^0)) \end{aligned} \quad (4.49)$$

pois $\hat{y}_k^0 \xrightarrow{K} +\infty$.

(a) Pelo Corolário 4.10, temos

$$\frac{1}{\mu_{k-1}^R} (z_{k-1}^0 + (\nabla G(x_{k-1})^T H(x_{k-1}) + \nabla H(x_{k-1})^T G(x_{k-1}))^T (x_k - x_{k-1}) + (z_k^0 - z_{k-1}^0)) \rightarrow 0. \quad (4.50)$$

Por outro lado, como os iterandos são de tipo O e V, temos pela condição (3.23) que $\{y_{k-1}^{E,0}\}$ é limitada. Logo multiplicando por μ_{k-1}^R a equação (4.49) e tendo em conta (4.50) temos

$$\hat{y}_k^0 \mu_{k-1}^R \xrightarrow{K} 0.$$

Considerando $k \in K$ suficientemente grande, dividindo (4.49) por \hat{y}_k^0 e considerando (4.50) temos

$$\frac{1}{\hat{y}_k^0 \mu_{k-1}^R} G(x_{k-1})^T H(x_{k-1}) \xrightarrow{K} 1, \quad (4.51)$$

de onde, pela positividade de \hat{y}_k^0 e μ_{k-1}^R , e a parte (a) segue que

$$G(x_{k-1})^T H(x_{k-1}) > 0 \quad (4.52)$$

para todo $k \in K$ suficientemente grande.

(b) Se $I_+ \neq \emptyset$, então $H_{I_+}(x^*) > 0$, logo pelo Corolário 4.5 (2), $[\hat{y}_k^H]_{I_+} = 0$ para $k \in K$ suficientemente grande. Assim em (4.37) temos que

$$\frac{\nu_{I_+}^k}{\hat{y}_k^0} = -G_{I_+}(x_k) \xrightarrow{K} 0. \quad (4.53)$$

Por outro lado, de (4.48), temos

$$\frac{1}{\hat{y}_k^0} \nabla f(x_k) - \nabla G(x_k)^T \frac{\gamma^k}{\hat{y}_k^0} - \nabla H_{I_0}(x_k)^T \frac{\nu_{I_0}^k}{\hat{y}_k^0} - \sum_{i \in \mathcal{A}(x^*)} \frac{[w_k^x]_i}{\hat{y}_k^0} e_i = \frac{\tilde{\varepsilon}_k}{\hat{y}_k^0} + \nabla H_{I_+}(x_k)^T \frac{\nu_{I_+}^k}{\hat{y}_k^0}$$

onde, aplicando o Lema 4.11 com $\alpha_k = \frac{1}{\hat{y}_k^0}$ e $\beta_k = \frac{\tilde{\varepsilon}_k}{\hat{y}_k^0} + (\nabla H_{I_+}(x_k))^T \frac{\nu_{I_+}^k}{\hat{y}_k^0}$ e considerando (4.53), concluimos que

$$\frac{\gamma^k}{\hat{y}_k^0} \xrightarrow{K} 0 \quad \text{e} \quad \frac{\nu_{I_0}^k}{\hat{y}_k^0} \xrightarrow{K} 0.$$

(c) Se $I_+ = \emptyset$ então $I_0 = \{1, \dots, l\} \neq \emptyset$. Logo suponhamos que $I_+ \neq \emptyset$. Note que de (4.37)

$$\frac{\gamma_{I_+}^k}{\hat{y}_k^0} = \frac{[\hat{y}_k^G]_{I_+}}{\hat{y}_k^0} - H_{I_+}(x_k)$$

e como $H_{I_+}(x^*) > 0$, então pela parte (b) temos que $[\hat{y}_k^G]_i \rightarrow +\infty$ para todo $i \in I_+$. Logo $[\hat{y}_k^G]_i$ satisfaz (4.31) para todo $k \in K$ suficientemente grande, e assim em (4.37), temos, para todo $k \in K$ suficientemente grande,

$$\gamma_{I_+}^k = -[y_{k-1}^{E,G}]_{I_+} - \frac{1}{\mu_{k-1}^R} G_{I_+}(x_{k-1}) + [F_k^{G_i}]_{I_+} - \hat{y}_k^0 H_{I_+}(x_k),$$

de onde segue que

$$\frac{\gamma_{I_+}^k}{\hat{y}_k^0} = -\frac{[y_{k-1}^{E,G}]_{I_+}}{\hat{y}_k^0} - \frac{1}{\hat{y}_k^0 \mu_{k-1}^R} G_{I_+}(x_{k-1}) + \frac{[F_k^{G_i}]_{I_+}}{\hat{y}_k^0} - H_{I_+}(x_k).$$

Pela parte (b) e (4.34) temos

$$\frac{1}{\hat{y}_k^0 \mu_{k-1}^R} G_{I_+}(x_{k-1}) + H_{I_+}(x_k) \xrightarrow{K} 0 \quad (4.54)$$

e assim

$$\frac{1}{\hat{y}_k^0 \mu_{k-1}^R} G_{I_+}(x_{k-1})^T H_{I_+}(x_{k-1}) \xrightarrow{K} -\|H_{I_+}(x^*)\|^2 < 0, \quad (4.55)$$

de onde, pela positividade de \hat{y}_k^0 e μ_{k-1}^R , e a parte (a) segue que

$$G_{I_+}(x_{k-1})^T H_{I_+}(x_{k-1}) < 0 \quad (4.56)$$

para todo $k \in K$ suficientemente grande. Assim por (4.52) temos $I_0 \neq \emptyset$.

(d) Se $I_+ = \emptyset$ então o resultado segue de (4.51). Caso contrario, se $I_+ \neq \emptyset$, então como

$$\frac{1}{\hat{y}_k^0 \mu_{k-1}^R} G_{I_0}(x_{k-1})^T H_{I_0}(x_{k-1}) + \frac{1}{\hat{y}_k^0 \mu_{k-1}^R} G_{I_+}(x_{k-1})^T H_{I_+}(x_{k-1}) = \frac{1}{\hat{y}_k^0 \mu_{k-1}^R} G(x_{k-1})^T H(x_{k-1})$$

temos de (4.51) e (4.55)

$$\frac{1}{\hat{y}_k^0 \mu_{k-1}^R} G_{I_0}(x_{k-1})^T H_{I_0}(x_{k-1}) \rightarrow \|H_{I_+}(x^*)\|^2 + 1 = M > 0.$$

(e) Seja $i \in I_+$. Como

$$\frac{1}{\hat{y}_k^0 \mu_{k-1}^R} G_i(x_k) + H_i(x_k) = \frac{1}{\hat{y}_k^0 \mu_{k-1}^R} G_i(x_{k-1}) + H_i(x_k) + \frac{1}{\hat{y}_k^0} \left[\frac{1}{\mu_{k-1}^R} (G_i(x_k) - G_i(x_{k-1})) \right]$$

temos pelo Corolário 4.10 e (4.54)

$$\frac{1}{\hat{y}_k^0 \mu_{k-1}^R} G_{I_+}(x_k) + H_{I_+}(x_k) \xrightarrow{K} 0$$

de onde, considerando (4.53), temos

$$\frac{\nu_{I_+}^k}{(\hat{y}_k^0)^2 \mu_{k-1}^R} \xrightarrow{K} H_{I_+}(x^*).$$

□

Dadas duas relações definidas em $\tilde{K} \times J$ com \tilde{K} infinito e J finito, o seguinte lema permite particionar J em dois subconjuntos satisfazendo uma propriedade que será útil em nossa análise.

Lema 4.15. *Sejam $\tilde{K} \subseteq \mathbb{N}$ um conjunto infinito, $J \subseteq \mathbb{N}$ um conjunto finito e $P_1(k, i)$ e $P_2(k, i)$ duas relações definidas em $\tilde{K} \times J$. Existe um conjunto infinito $K_l \subseteq \tilde{K}$ e conjuntos disjuntos J_0 e J_+ com $J = J_0 \cup J_+$ onde J_+ está formado por todos os $i \in J$ tais que $P_1(k, i)$ e $P_2(k, i)$ são satisfeitas simultaneamente somente para um subconjunto finito de K_l .*

Demonstração. Definimos primeiramente $J_0 = \emptyset$ e $K_1 = \tilde{K}$ e $J_+ = J \setminus J_0$. Se para algum $i \in J_+$ existe um conjunto infinito $K_2 \subseteq K_1$ para o qual vale $P_1(k, i)$ ou $P_2(k, i)$ para todo $k \in K_2$ adicionamos i a J_0 e repetimos o processo considerando agora K_2 no lugar de K_1 . Como J é finito conseguimos conjuntos J_0 , J_+ e K_l onde J_+ está formado por todos os

$i \in J$ tais que $P_1(k, i)$ e $P_2(k, i)$ são satisfeitas somente para um subconjunto finito de K_l .

□

Note que, dados dois índices i_1 e i_2 pode acontecer que existam conjuntos infinitos disjuntos K_1 e \tilde{K}_2 tais que, vale $P_1(k, i_1)$ ou $P_2(k, i_1)$ para todo $k \in K_1$, e vale $P_1(k, i_2)$ ou $P_2(k, i_2)$ para todo $k \in \tilde{K}_2$, porém neste caso \tilde{K}_2 não é considerada, pois por construção $K_2 \subseteq K_1$.

Como todo ponto M-estacionário precisa ser fracamente estacionário, no seguinte resultado mostramos que se x^* um ponto de acumulação viável para o Problema (4.1) da sequência $\{x_k\}_S$ gerada pelo Algoritmo 5, então x^* é fracamente estacionário. Isto nos permite definir os multiplicadores ν^* , γ^* e w^* para a análise de M-estacionaridade do Problema (4.1).

Teorema 4.16. *Seja x^* um ponto de acumulação viável para o Problema (4.1) da sequência $\{x_k\}_S$ gerada pelo Algoritmo 5. Se x^* satisfaz as condições de MPEC-LICQ, então existem um conjunto infinito $K_0 \subseteq K$, $\nu^*, \gamma^* \in \mathbb{R}^s$ e $w^* \in \mathbb{R}^n$ tais que*

$$\nu^k \xrightarrow{K_0} \nu^*, \quad \gamma^k \xrightarrow{K_0} \gamma^* \quad e \quad w_k^x \xrightarrow{K_0} w^* \quad (4.57)$$

com $\nu_{I_+}^* = 0$, $w_i^* = 0$ se $i \in \mathcal{F}(x^*)$ e

$$\nabla f(x^*) - \nabla G(x^*)^T \gamma^* - \nabla H_{I_0}(x^*)^T \nu_{I_0}^* - w^* = 0,$$

isto é, x^* é fracamente estacionário.

Demonstração. Seja $x_k \xrightarrow{K} x^*$ para $K \subseteq \mathcal{S}$. Note que podemos assumir que $\hat{y}_k^0 \xrightarrow{K} +\infty$, caso contrário pelo analisado no início da Seção 4.4.2, teríamos que x^* é S-estacionário. Se $I_+ = \emptyset$ de (4.38) temos que

$$\nabla f(x_k) - \nabla G(x_k)^T \gamma^k - \nabla H_{I_0}(x_k)^T \nu_{I_0}^k - \sum_{i \in \mathcal{A}(x^*)} [w_k^x]_i e_i = \tilde{\varepsilon}_k$$

com $\tilde{\varepsilon}_k \rightarrow 0$. Logo podemos aplicar diretamente o Lema 4.11 com $\alpha_k = 1$ e $\beta_k = \tilde{\varepsilon}_k$ para obter o resultado. Se $I_+ \neq \emptyset$, suponhamos que para algum $i \in I_0$ exista um conjunto infinito $K_i \subseteq K$ tal que a condição (4.30) é satisfeita para todo $k \in K_i$; neste caso, substituindo (4.31) em (4.37) para $k \in K_i$ suficientemente grande, resulta que

$$\gamma_i^k = -[y_{k-1}^{E,G}]_i - \frac{1}{\mu_{k-1}^R} G_i(x_{k-1}) + F_k^{G_i} - \hat{y}_k^0 H_i(x_k).$$

Como os iterandos são de tipo O e V, temos pela condição (3.23) que $\{y_{k-1}^{E,G}\}$ é limitada, além disso por (4.34) $F_k^{G_i} \rightarrow 0$. Logo, aplicando o Lema 4.14 (b) na última equação, temos

$$\frac{1}{\hat{y}_k^0 \mu_{k-1}^R} G_i(x_{k-1}) \xrightarrow{K_i} 0$$

de onde

$$\frac{1}{\hat{y}_k^0 \mu_{k-1}^R} G_i(x_{k-1}) H_i(x_{k-1}) \xrightarrow{K_i} 0. \quad (4.58)$$

De maneira similar, (4.58) vale se considerarmos que para algum $i \in I_0$ exista um conjunto infinito K_i tal que a condição (4.32) é satisfeita para todo $k \in K_i$.

Pelo Lema 4.15 para $\tilde{K} = K$ e $J = I_0$, existem conjuntos J_0 e K_l tais que para todo $i \in J_+ = I_0 \setminus J_0$ as condições (4.30) e (4.32) valem somente para um conjunto finito de pontos x_k com $k \in K_l$. Como $I_0 = J_+ \cup J_0$ e $K_l \subseteq K$, temos do Lema 4.14 (d) que

$$\frac{1}{\hat{y}_k^0 \mu_{k-1}^R} G_{J_0}(x_{k-1})^T H_{J_0}(x_{k-1}) + \frac{1}{\hat{y}_k^0 \mu_{k-1}^R} G_{J_+}(x_{k-1})^T H_{J_+}(x_{k-1}) \xrightarrow{K_l} M > 0. \quad (4.59)$$

Pela construção de K_l , temos que $K_l \subseteq K_{l-1} \subseteq \dots \subseteq K_1 = K$. Logo como a condição (4.58) vale para todo $i \in J_0$, temos que

$$\frac{1}{\hat{y}_k^0 \mu_{k-1}^R} G_{J_0}(x_{k-1})^T H_{J_0}(x_{k-1}) \xrightarrow{K_l} 0$$

Assim, em (4.59), $J_+ \neq \emptyset$ e

$$\frac{1}{\hat{y}_k^0 \mu_{k-1}^R} G_{J_+}(x_{k-1})^T H_{J_+}(x_{k-1}) \xrightarrow{K_l} M.$$

Por outro lado, se $i \in J_+$, então

$$\begin{aligned} \frac{1}{\hat{y}_k^0 \mu_{k-1}^R} G_i(x_k) H_i(x_k) &= \frac{1}{\hat{y}_k^0 \mu_{k-1}^R} [G_i(x_{k-1}) H_i(x_{k-1}) + G_i(x_k) (H_i(x_k) - H_i(x_{k-1})) \\ &\quad + (G_i(x_k) - G_i(x_{k-1})) H_i(x_{k-1})] \\ &= \frac{1}{\hat{y}_k^0 \mu_{k-1}^R} G_i(x_{k-1}) H_i(x_{k-1}) + \frac{1}{\hat{y}_k^0} G_i(x_k) \left[\frac{1}{\mu_{k-1}^R} (H_i(x_k) - H_i(x_{k-1})) \right] \\ &\quad + \frac{1}{\hat{y}_k^0} H_i(x_{k-1}) \left[\frac{1}{\mu_{k-1}^R} (G_i(x_k) - G_i(x_{k-1})) \right], \end{aligned}$$

logo, pelo Corolário 4.10,

$$\lim_{k \in K_l} \frac{1}{\hat{y}_k^0 \mu_{k-1}^R} G_i(x_k) H_i(x_k) = \lim_{k \in K_l} \frac{1}{\hat{y}_k^0 \mu_{k-1}^R} G_i(x_{k-1}) H_i(x_{k-1}).$$

Assim

$$\frac{1}{\hat{y}_k^0 \mu_{k-1}^R} G_{J_+}(x_k)^T H_{J_+}(x_k) \xrightarrow{K_l} M. \quad (4.60)$$

Agora, posto que para $i \in J_+$ (4.30) e (4.32) valem somente para um conjunto finito de pontos x_k com $k \in K_l$ temos de (4.15), (4.21) e (4.37) que para todo $k \in K_l$ suficientemente

grande

$$\begin{aligned}\gamma_{J_+}^k &= -\hat{y}_k^0 H_{J_+}(x_k) \\ \nu_{J_+}^k &= -\hat{y}_k^0 G_{J_+}(x_k),\end{aligned}$$

logo em (4.60)

$$\frac{1}{(\hat{y}_k^0)^3 \mu_{k-1}^R} (\gamma_{J_+}^k)^T \nu_{J_+}^k \xrightarrow{K_l} M. \quad (4.61)$$

Afirmamos que $I_+ = \emptyset$ ou

$$\nu_{I_+}^k \xrightarrow{K_l} 0.$$

De fato, suponhamos que $I_+ \neq \emptyset$ e que $\nu_{I_+}^k$ não convirja para 0 em K_l , então existe $\varepsilon > 0$ e um conjunto infinito $K_l^0 \subseteq K_l$ tal que

$$\|\nu_{I_+}^k\| \geq \varepsilon \quad (4.62)$$

para todo $k \in K_l^0$. Temos do Lema 4.14 (e) que

$$\frac{\|\nu_{I_+}^k\|^2}{(\hat{y}_k^0)^4 (\mu_{k-1}^R)^2} \xrightarrow{K_l^0} \|H_{I_+}(x^*)\|^2$$

e combinando este resultado com (4.61) resulta

$$\frac{y_k^0 \mu_{k-1}^R}{\|\nu_{I_+}^k\|^2} (\gamma_{J_+}^k)^T \nu_{J_+}^k \xrightarrow{K_l^0} \frac{M}{\|H_{I_+}(x^*)\|^2}.$$

Assim pelo Lema 4.14 (a) podemos concluir que

$$\frac{(\gamma_{J_+}^k)^T \nu_{J_+}^k}{\|\nu_{I_+}^k\|^2} \xrightarrow{K_l^0} \infty$$

e como

$$(\gamma_{J_+}^k)^T \nu_{J_+}^k \leq \|(\gamma_{J_+}^k)\| \|\nu_{J_+}^k\| \leq \|(\gamma_{J_+}^k, \nu_{J_+}^k)\|^2 \leq \|(\gamma^k, \nu_{I_0}^k)\|^2,$$

então

$$\frac{\|(\gamma^k, \nu_{I_0}^k)\|}{\|\nu_{I_+}^k\|} \xrightarrow{K_l^0} +\infty$$

e por (4.62), devemos ter

$$\|(\gamma^k, \nu_{I_0}^k)\| \xrightarrow{K_l^0} +\infty.$$

Concluimos dos últimos limites que

$$\frac{\nu_{I_+}^k}{\|(\gamma^k, \nu_{I_0}^k)\|} \xrightarrow{K_l^0} 0 \quad \text{e} \quad \frac{1}{\|(\gamma^k, \nu_{I_0}^k)\|} \xrightarrow{K_l^0} 0. \quad (4.63)$$

Dividindo (4.48) por $\|(\gamma^k, \nu_{I_0}^k)\|$ temos

$$\begin{aligned} & \frac{1}{\|(\gamma^k, \nu_{I_0}^k)\|} \left(\nabla f(x_k) - \nabla G(x_k)^T \gamma^k - \nabla H_{I_0}(x_k)^T \nu_{I_0}^k - \sum_{i \in \mathcal{A}(x^*)} [w_k^x]_i e_i \right) \\ &= \nabla H_{I_+}(x_k)^T \frac{\nu_{I_+}^k}{\|(\gamma^k, \nu_{I_0}^k)\|} + \frac{\varepsilon_k}{\|(\gamma^k, \nu_{I_0}^k)\|} \end{aligned}$$

onde, considerando (4.63), podemos aplicar o Lema 4.11 com

$$\alpha_k = \frac{1}{\|(\gamma^k, \nu_{I_0}^k)\|} \quad \text{e} \quad \beta_k = \nabla H_{I_+}(x_k)^T \frac{\nu_{I_+}^k}{\|(\gamma^k, \nu_{I_0}^k)\|} + \frac{\varepsilon_k}{\|(\gamma^k, \nu_{I_0}^k)\|},$$

resultando que

$$\frac{(\gamma^k, \nu_{I_0}^k)}{\|(\gamma^k, \nu_{I_0}^k)\|} \xrightarrow{K_l^0} 0,$$

o que é impossível. Assim, devemos ter

$$\nu_{I_+}^k \xrightarrow{K_l} 0.$$

Agora, de (4.48), temos

$$\nabla f(x_k) - \nabla G(x_k)^T \gamma^k - \nabla H_{I_0}(x_k)^T \nu_{I_0}^k - \sum_{i \in \mathcal{A}(x^*)} [w_k^x]_i e_i = \nabla H_{I_+}(x_k)^T \nu_{I_+}^k + \tilde{\varepsilon}_k,$$

e então aplicando o Lema 4.11 para $k \in K_l$ com $\alpha_k = 1$ e $\beta_k = \nabla H_{I_+}(x_k)^T \nu_{I_+}^k + \tilde{\varepsilon}_k$, concluimos que existem γ^* , $\nu_{I_0}^*$ e w_i^* para $i \in \mathcal{A}(x^*)$ tais que

$$\gamma^k \xrightarrow{K_l} \gamma^*, \quad \nu_{I_0}^k \xrightarrow{K_l} \nu_{I_0}^* \quad \text{e} \quad [w_k^x]_i \xrightarrow{K_l} w_i^*.$$

Finalmente para obter (4.57) consideramos $\nu_{I_+}^* = 0$, $w_i^* = 0$ se $i \in \mathcal{F}(x^*)$ e $K_0 = K_l$. \square

A seguir apresentamos outro resultado, dentre os principais desta tese. Neste, mostramos convergência para pontos M -estacionários, quando o Algoritmo 5 é aplicado ao Problema MPEC (4.1). Este resultado é equivalente ao que foi estabelecido em [1] para métodos baseados em Lagrangiano Aumentado e assim como em [1], exigiremos válida a condição MPEC-LICQ.

Teorema 4.17. *Seja x^* um ponto de acumulação viável para o Problema (4.1) da sequência $\{x_k\}_S$ gerada pelo Algoritmo 5. Se x^* satisfaz a condição MPEC-LICQ, então x^* é pelo*

menos M -estacionário.

Demonstração. Seja $x_k \xrightarrow{K} x^*$ para $K \subseteq \mathcal{S}$. Pelo Teorema 4.16, temos que x^* é fracamente estacionário. Considerando os multiplicadores ν^*, γ^* e w^* do Teorema 4.16 onde $\gamma^k \xrightarrow{K_0} \gamma^*$ e $\nu^k \xrightarrow{K_0} \nu^*$, falta somente demonstrar que $(\gamma_i^* > 0 \text{ e } \nu_i^* > 0)$ ou $\gamma_i^* \nu_i^* = 0$ para todo $i \in I_0$. De (4.37) temos

$$\begin{aligned}\gamma_i^k &= [\hat{y}_k^G]_i - \hat{y}_k^0 H_i(x_k) \text{ e} \\ \nu_i^k &= [\hat{y}_k^H]_i - \hat{y}_k^0 G_i(x_k).\end{aligned}$$

Temos três possibilidades para $i \in I_0$: que existam conjunto infinitos $K_1, K_2, K_3 \subseteq K_0$ tais que

1. $[\hat{y}_k^G]_i > 0$ e $[\hat{y}_k^H]_i > 0$ para todo $k \in K_1$,
2. $([\hat{y}_k^G]_i = 0 \text{ e } [\hat{y}_k^H]_i > 0)$ ou $([\hat{y}_k^G]_i > 0 \text{ e } [\hat{y}_k^H]_i = 0)$ para todo $k \in K_2$ e
3. $[\hat{y}_k^G]_i = 0$ e $[\hat{y}_k^H]_i = 0$ para todo $k \in K_3$.

Por (4.15) e (4.21), $[\hat{y}_k^G]_i > 0$ equivale a (4.30) para todo k suficientemente grande, e $[\hat{y}_k^H]_i > 0$ equivale a (4.32) para todo k suficientemente grande. Pelo Lema 4.15 considerando $\tilde{K} = K_0$ e $J = I_0$, temos que existem conjuntos J_0, J_+ e K_l ($K_l \subseteq K_0$) tais que para todo $i \in J_+ = I_0 \setminus J_0$ as condições (4.30) e (4.32) valem somente para um conjunto finito de pontos x_k com $k \in K_l$. Fazendo a mesma análise que na demonstração do Teorema 4.16 temos que a condição (4.61) é válida para os novos J_0, J_+ e K_l , isto é,

$$\frac{1}{(\hat{y}_k^0)^3 \mu_{k-1}^R} (\gamma_{J_+}^k)^T \nu_{J_+}^k \xrightarrow{K_l} M. \quad (4.64)$$

Pelo Teorema 4.16, as sequências $\{\gamma_k\}$ e $\{\nu_k\}$ são convergentes em K_0 , e $\hat{y}_k^0 \xrightarrow{K} \infty$, então

$$\frac{(\gamma_{J_+}^k)^T \nu_{J_+}^k}{\hat{y}_k^0} \xrightarrow{K_0} 0.$$

Logo por (4.64)

$$\frac{1}{(\hat{y}_k^0)^2 \mu_{k-1}^R} \xrightarrow{K_l} \infty. \quad (4.65)$$

Por outro lado, como $I_0 = J_0 \cup J_+$, devemos mostrar que $(\gamma_i^* > 0 \text{ e } \nu_i^* > 0)$ ou $\gamma_i^* \nu_i^* = 0$ para $i \in J_0$ e $i \in J_+$. Logo as três possibilidades anteriores são equivalentes a:

1. $i \in J_0$ e exista $K_l^1 \subseteq K_l$ tal que as condições (4.30) e (4.32) se satisfazem para todo $k \in K_l^1$ suficientemente grande.
2. $i \in J_0$ e exista $K_l^2 \subseteq K_l$ tal que somente uma das condições (4.30) ou (4.32) se satisfaz para todo $k \in K_l^2$ suficientemente grande.

3. $i \in J_+$ e exista $K_l^3 \subseteq K_l$ tal que as condições (4.30) e (4.32) não se satisfazem para todo $k \in K_l^3$ suficientemente grande.

Caso 1: Se $i \in J_0$ e existe $K_l^1 \subseteq K_l$ tal que (4.30) e (4.32) se satisfazem para todo $k \in K_l^1$. Considerando $k \in K_l^1$ suficientemente grande e substituindo (4.31) e (4.33) em (4.37), resulta que

$$\gamma_i^k = -[y_{k-1}^{E,G}]_i - \frac{1}{\mu_{k-1}^R} G_i(x_{k-1}) + F_k^{G_i} - \hat{y}_k^0 H_i(x_k), \quad (4.66)$$

$$\nu_i^k = -[y_{k-1}^{E,H}]_i - \frac{1}{\mu_{k-1}^R} H_i(x_{k-1}) + F_k^{H_i} - \hat{y}_k^0 G_i(x_k). \quad (4.67)$$

Multiplicando (4.66) por \hat{y}_k^0 , (4.67) por $\frac{-1}{\mu_{k-1}^R}$ e somando os lados das equações resultantes, obtemos

$$\begin{aligned} \hat{y}_k^0 \gamma_i^k - \frac{\nu_i^k}{\mu_{k-1}^R} &= -\hat{y}_k^0 [y_{k-1}^{E,G}]_i - \frac{\hat{y}_k^0}{\mu_{k-1}^R} G_i(x_{k-1}) + \hat{y}_k^0 F_k^{G_i} - (\hat{y}_k^0)^2 H_i(x_k) \\ &\quad + \frac{1}{\mu_{k-1}^R} [y_{k-1}^{E,H}]_i + \frac{1}{(\mu_{k-1}^R)^2} H_i(x_{k-1}) - \frac{1}{\mu_{k-1}^R} F_k^{H_i} + \frac{\hat{y}_k^0}{\mu_{k-1}^R} G_i(x_k) \\ &= -\hat{y}_k^0 [y_{k-1}^{E,G}]_i + \frac{1}{\mu_{k-1}^R} [y_{k-1}^{E,H}]_i + \frac{\hat{y}_k^0}{\mu_{k-1}^R} (G_i(x_k) - G_i(x_{k-1})) \\ &\quad + \hat{y}_k^0 F_k^{G_i} - \frac{1}{\mu_{k-1}^R} F_k^{H_i} + \frac{1}{(\mu_{k-1}^R)^2} H_i(x_{k-1}) - (\hat{y}_k^0)^2 H_i(x_k) \\ &= -\hat{y}_k^0 [y_{k-1}^{E,G}]_i + \frac{1}{\mu_{k-1}^R} [y_{k-1}^{E,H}]_i + \frac{\hat{y}_k^0}{\mu_{k-1}^R} (G_i(x_k) - G_i(x_{k-1})) \\ &\quad + \hat{y}_k^0 F_k^{G_i} - \frac{1}{\mu_{k-1}^R} F_k^{H_i} - \frac{1}{(\mu_{k-1}^R)^2} (H_i(x_k) - H_i(x_{k-1})) \\ &\quad + \left[\frac{1}{(\mu_{k-1}^R)^2} - (\hat{y}_k^0)^2 \right] H_i(x_k), \end{aligned}$$

de onde segue que

$$\begin{aligned} \left[(\hat{y}_k^0)^2 - \frac{1}{(\mu_{k-1}^R)^2} \right] H_i(x_k) &= -\hat{y}_k^0 [y_{k-1}^{E,G}]_i + \frac{1}{\mu_{k-1}^R} [y_{k-1}^{E,H}]_i + \frac{\hat{y}_k^0}{\mu_{k-1}^R} (G_i(x_k) - G_i(x_{k-1})) \\ &\quad + \hat{y}_k^0 F_k^{G_i} - \frac{1}{\mu_{k-1}^R} F_k^{H_i} - \frac{1}{(\mu_{k-1}^R)^2} (H_i(x_k) - H_i(x_{k-1})) \\ &\quad - \hat{y}_k^0 \gamma_i^k + \frac{\nu_i^k}{\mu_{k-1}^R} \\ &= \frac{1}{\mu_{k-1}^R} \left[-\hat{y}_k^0 \mu_{k-1}^R [y_{k-1}^{E,G}]_i + [y_{k-1}^{E,H}]_i + \hat{y}_k^0 \mu_{k-1}^R F_k^{G_i} - F_k^{H_i} \right. \\ &\quad + \hat{y}_k^0 \mu_{k-1}^R \left(\frac{1}{\mu_{k-1}^R} (G_i(x_k) - G_i(x_{k-1})) \right) - \hat{y}_k^0 \mu_{k-1}^R \gamma_i^k + \nu_i^k \\ &\quad \left. - \frac{1}{\mu_{k-1}^R} (H_i(x_k) - H_i(x_{k-1})) \right]. \end{aligned}$$

Assim

$$\begin{aligned}
((\hat{y}_k^0 \mu_{k-1}^R)^2 - 1) \frac{H_i(x_k)}{\mu_{k-1}^R} &= -\hat{y}_k^0 \mu_{k-1}^R [y_{k-1}^{E,G}]_i + [y_{k-1}^{E,H}]_i + \hat{y}_k^0 \mu_{k-1}^R F_k^{G_i} - F_k^{H_i} \\
&\quad + \hat{y}_k^0 \mu_{k-1}^R \left(\frac{1}{\mu_{k-1}^R} (G_i(x_k) - G_i(x_{k-1})) \right) - \hat{y}_k^0 \mu_{k-1}^R \gamma_i^k + \nu_i^k \\
&\quad - \frac{1}{\mu_{k-1}^R} (H_i(x_k) - H_i(x_{k-1})). \tag{4.68}
\end{aligned}$$

Como os iterandos são de tipo O e V, temos pela condição (3.23) que $\{y_{k-1}^{E,G}\}$ e $\{y_{k-1}^{E,H}\}$ são limitadas. Além disso, neste caso $\{\gamma_k\}$ e $\{\nu_k\}$ são convergentes em $K_l^1 \subseteq K$. Logo pelo Lema 4.14 (a) e pelo Corolário 4.10 o lado direito de (4.68) é limitado em K_l^1 . Em consequência de (4.68), temos que

$$\left\{ \frac{H_i(x_k)}{\mu_{k-1}^R} \right\} \text{ é limitado em } K_l^1$$

e assim, pelo Lema 4.14 (a),

$$\hat{y}_k^0 H_i(x_k) = \hat{y}_k^0 \mu_{k-1}^R \frac{H_i(x_k)}{\mu_{k-1}^R} \xrightarrow{K_l^1} 0. \tag{4.69}$$

Repetindo os mesmos argumentos para $G_i(x_k)$ podemos mostrar também que

$$\hat{y}_k^0 G_i(x_k) \xrightarrow{K_l^1} 0. \tag{4.70}$$

Considerando (4.69) e (4.70) em (4.37), temos que $\gamma_i^* \geq 0$ e $\nu_i^* \geq 0$.

Caso 2: Suponhamos $i \in J_0$ e que exista $K_l^2 \subseteq K_l$ tal que somente uma das condições (4.30) ou (4.32) se satisfaz para todo $k \in K_l^2$. Assumamos sem perda de generalidade que (4.30) é válida. Considerando $k \in K_l^2$ suficientemente grande e substituindo (4.31) em (4.37) temos que (4.66) é válida para todo $k \in K_l^2$, isto é,

$$\gamma_i^k = -[y_{k-1}^{E,G}]_i - \frac{1}{\mu_{k-1}^R} G_i(x_{k-1}) + F_k^G - \hat{y}_k^0 H_i(x_k) \tag{4.71}$$

e como (4.32) não é válida para todo $k \in K_l^2$, temos de (4.15) e (4.37) que

$$\nu_i^k = -\hat{y}_k^0 G_i(x_k). \tag{4.72}$$

Afirmamos que $\nu_i^* = 0$. De fato se $\nu_i^* \neq 0$, temos de (4.71) que

$$\gamma_i^k = -[y_{k-1}^{E,G}]_i - \frac{1}{\mu_{k-1}^R} (G_i(x_{k-1}) - G_i(x_k) + G_i(x_k)) + F_k^G - \hat{y}_k^0 H_i(x_k),$$

e então por (4.72)

$$\gamma_i^k = -[y_{k-1}^{E,G}]_i - \frac{1}{\mu_{k-1}^R} (G_i(x_{k-1}) - G_i(x_k)) + \frac{1}{\mu_{k-1}^R} \frac{\nu_i^k}{\hat{y}_k^0} + F_k^G - \hat{y}_k^0 H_i(x_k).$$

de onde segue que

$$\frac{\nu_i^k}{(\hat{y}_k^0)^2 \mu_{k-1}^R} = \frac{[y_{k-1}^{E,G}]_i + \gamma_i^k}{\hat{y}_k^0} - \frac{1}{\hat{y}_k^0} \left[\frac{1}{\mu_{k-1}^R} (G_i(x_k) - G_i(x_{k-1})) + F_k^G \right] + H_i(x_k).$$

Como no caso anterior $\{y_{k-1}^{E,G}\}$ é limitada. Logo como $G_i(x^*) = 0$ e $\{\gamma_i^k\}$ é limitada, temos pelo Corolário 4.10 e (4.34) que

$$\frac{\nu_i^k}{(\hat{y}_k^0)^2 \mu_{k-1}^R} \xrightarrow{K^2} 0.$$

Porém como $\nu_i^* \neq 0$ então a última condição contradiz (4.65).

Caso 3: Se $i \in J_+$ e existe $K_l^3 \subseteq K_l$ tal que (4.30) e (4.32) não se satisfazem para nenhum $k \in K_l^3$. Como as condições (4.30) e (4.32) não são válidas para nenhum $k \in K_l^3$, temos por (4.15) que para todo $k \in K_l^3$ suficientemente grande $[\hat{y}_k^H]_i = 0$ e $[\hat{y}_k^G]_i = 0$. Logo em (4.37)

$$\nu_i^k = -\hat{y}_k^0 G_i(x_k) \text{ e } \gamma_i^k = -\hat{y}_k^0 H_i(x_k)$$

de onde

$$\nu_i^k = -\hat{y}_k^0 \mu_{k-1}^R \frac{G_i(x_k)}{\mu_{k-1}^R}. \quad (4.73)$$

Afirmamos que $\nu_i^* = 0$ e $\gamma_i^* = 0$. De fato se $\nu_i^* \neq 0$, então por (4.73) e o Lema 4.14 (a),

$$\left| \frac{G_i(x_k)}{\mu_{k-1}^R} \right| \xrightarrow{K^3} \infty, \quad (4.74)$$

e como

$$\frac{G_i(x_{k-1})}{\mu_{k-1}^R} = \frac{G_i(x_k)}{\mu_{k-1}^R} - \frac{G_i(x_k) - G_i(x_{k-1})}{\mu_{k-1}^R} \quad (4.75)$$

temos pelo Corolário 4.10 e a relação (4.74) que

$$\left| \frac{G_i(x_{k-1})}{\mu_{k-1}^R} \right| \xrightarrow{K^3} \infty. \quad (4.76)$$

Como (4.30) não é válida para todo $k \in K_l^3$, então

$$[y_{k-1}^{E,G}]_i + \frac{1}{\mu_{k-1}^R} G_i(x_{k-1}) + F_k^{G_i} \geq -\varepsilon_k,$$

Nesta expressão ao igual que nos casos anteriores $\{y_{k-1}^{E,G}\}$ é limitada, logo considerando

(4.34) e a relação (4.76) temos

$$\frac{G_i(x_{k-1})}{\mu_{k-1}^R} \xrightarrow{K^3} \infty. \quad (4.77)$$

Finalmente, considerando $g_j(x) = -G_i(x)$ em (4.21) temos para todo $k \in K_l^3$ suficientemente grande

$$[y_k]_j = [y_{k-1}^{E,G}]_i + \frac{1}{\mu_{k-1}^R} G_i(x_{k-1}) - F_k^{G_i}$$

e assim por (4.34) e (4.77), $[y_k]_j \xrightarrow{K^3} \infty$, o que contradiz o fato dos multiplicadores $[y_k]_j$ serem limitados superiormente, pois pelo Corolário 4.5 (1), $[y_k]_j \leq \varepsilon_k$. Assim $\nu_i^* = 0$ e similarmente $\gamma_i^* = 0$. \square

4.5 Testes numéricos

Nesta seção, relatamos e discutimos o comportamento do Algoritmo 5. Os testes computacionais foram realizados usando 73 problemas (veja Tabela 4.1) da coleção MacMPEC que está disponível na linguagem de modelagem AMPL em <http://www.mcs.anl.gov/~leyffer/MacMPEC>. Por conta de dificuldades na interface entre AMPL e MATLAB, nos restringimos a problemas de dimensão menor o igual a 50. Além disso, omitimos os problemas que continham problemas de complementaridade mistos, multi-objetivo, inviáveis ou com variáveis inteiras. Os testes realizados servem para validar a nossa teoria destacando que em [1, 30, 40] os problemas MPEC da coleção MacMPEC foram satisfatoriamente resolvidos por outros métodos.

Os experimentos numéricos foram realizados usando uma implementação do Algoritmo 5 escrito em MATLAB R2015a, v. 8.5.0.197613, e foram feitos em um notebook ACER Intel Core i5-4210U, CPU 1.70GHz 2.40GHz, com 6GB RAM, Windows 8.1 Enterprise N 64-bits. Os valores dos parâmetros de controle e os valores iniciais requeridos pelo Algoritmo 5 foram considerados como em [30] e são especificados na seguinte tabela:

Parâmetro	Valor	Parâmetro	Valor	Parâmetro	Valor
$\phi_{V,0}^{\max}, \phi_{O,0}^{\max}$	$1.0e^{+3}$	μ_0^R	$1.0e^{-4}$	τ_s	$1.0e^{-5}$
ϵ_a	$1.0e^{-6}$	μ_0	1.0	β	$1.0e^{-5}$
$\bar{\gamma}$	0.5	γ_s	$1.0e^{-3}$	λ	0.2
y_{\max}	$1.0e^{+6}$	θ	$1.0e^{-5}$	τ_0	1.0

Inicializamos os multiplicadores iguais a 1. Quando não era fornecido o ponto inicial em geral foi tomado como sendo o vetor das entradas iguais a 1, com algumas exceções que mencionaremos na descrição da tabela. Além disso, se o algoritmo não convergir dentro de $k_{\max} = 500$ iterações, então foi considerado como tendo falhado. Os critérios de parada do algoritmo em uma solução aproximada ou um ponto estacionário da medida de inviabilidade são dadas por (3.39) e (3.41), respectivamente. Quando é

declarada uma solução aproximada é possível calcular os multiplicadores aproximados, usando as equações (4.15) e (4.16) o que permite, mediante as equações (4.37), a saber

$$\begin{aligned}\gamma^k &= \hat{y}_k^G - \hat{y}_k^0 H(x_k) \\ \nu^k &= \hat{y}_k^H - \hat{y}_k^0 G(x_k),\end{aligned}$$

analisar o sinal dos multiplicadores associados ao conjunto I_{00} . Nesse sentido, será declarado que o ponto encontrado é aproximadamente C-estacionário (-C) quando pelo menos um par de multiplicadores associados às restrições de complementaridade em I_{00} , parecem estar convergindo para valores negativos, isto é, se $\gamma_i^k < -1.0e^{-4}$ e $\nu_i^k < -1.0e^{-4}$ para algum $i \in I_{00}$. O ponto é aproximadamente M-estacionário (-M) quando se observa que existem multiplicadores associados às restrições complementaridade em I_{00} que são aproximadamente negativos, sendo necessariamente neste caso o multiplicador da restrição complementar associada aproximadamente zero, isto é, se existe $i \in I_{00}$ tal que $\gamma_i^k < -1.0e^{-4}$ e $|\gamma_i^k \nu_i^k| < 1.0e^{-4}$, ou $\nu_i^k < -1.0e^{-4}$ e $|\gamma_i^k \nu_i^k| < 1.0e^{-4}$. O ponto encontrado é aproximadamente S-estacionário (-S) quando todos os multiplicadores associados às restrições complementaridade em I_{00} são aproximadamente não-negativos, isto é, se $\gamma_i^k > -1.0e^{-4}$ e $\nu_i^k > -1.0e^{-4}$ para todo $i \in I_{00}$.

A Tabela 4.2 coleta os resultados obtidos com o Algoritmo 5. A primeira coluna “Problema” contém o nome do problema. A segunda coluna “Melh. Obj.” fornece o melhor valor de função objetivo obtido em <http://www.mcs.anl.gov/~leyffer/MacMPEC>. As próximas cinco colunas contêm informações sobre o desempenho do Algoritmo 5. A primeira dessas colunas “Obj.” contém o valor da função objetivo em x_{k_f} , o ponto ao qual o Algoritmo 5 convergiu na última iteração k_f . Em alguns casos quando este valor for maior que o melhor valor de função objetivo conhecido na literatura ou o algoritmo não convergiu, realizamos testes com diferentes pontos iniciais, selecionados aleatoriamente, fato que está simbolizado por (*) no final do valor. A segunda “ $r(x_{k_f}, y_{k_f})$ ” indica o valor da função proximal r em (x_{k_f}, y_{k_f}) , a qual é uma medida da viabilidade e otimalidade da solução x_{k_f} encontrada; em alguns casos quando o algoritmo não convergiu, porém o valor foi próximo de zero, foi aumentada a tolerância ($\tau_s = 1.0e^{-4}$), o que também é simbolizado por (*) nesta coluna. A terceira coluna “Inv.” indica a medida de inviabilidade [12] dada por

$$\max \left\{ \begin{array}{l} \|h(x_{k_f})\|^2, \|g(x_{k_f})_+\|^2, \|(-H(x_{k_f}))_+\|^2, \\ \|(-G(x_{k_f}))_+\|^2, \|(-x_{k_f})_+\|^2, \left((G(x_{k_f})^T H(x_{k_f}))_+ \right)^2 \end{array} \right\}.$$

A quarta coluna indica a ordem de convergência estimada (OCE) dada pela quantidade [31]

$$OCE = \log r(x_{k_f}, y_{k_f}) / \log r(x_{k_f-1}, y_{k_f-1}).$$

Para a análise da OCE consideramos ao igual que em [31] que a convergência é superlinear, se $OCE > 1.1$

A quinta coluna “Stat.” indica o status de ser aproximadamente C-, M- ou S-estacionário. Finalmente a última coluna fornece como foram a última ou as duas últimas direções de descida d , “1L” significa que a última foi local; “2L”, que as duas últimas foram locais; e “G”, que a última foi global. Esta análise é uma medida do tempo computacional. No caso de descida local, d se obtém como solução de um sistema de equações lineares, e no caso de descida global, d é solução de um problema quadrático (função quadprog.m do MATLAB).

Se o algoritmo falhou, isso é indicado por um traço em todas as seis colunas, indicando na última a razão pela qual o algoritmo não convergiu de acordo com Tabela 4.1.

Tabela 4.1: Razões pelas quais o algoritmo falhou

-1	Convergiu a uma solução inviável.
-2	Terminou com o número máximo de iterações. (iteraões de tipo F)
-3	Terminou com problemas numéricos.

Tabela 4.2: Resultados obtidos com o Algoritmo 5

Problema	Melh. Obj.	Obj.	$r(x_{k_f}, y_{k_f})$	Inv.	OCE	Stat.	d
bard1	17	17	4.6837e-08	2.1716e-15	2.3126	-S	1L
bard2	6598	6598	1.0461e-07	3.6325e-15	1.7961	-S	2L
bard3	-12.6787	-12.6787	3.4925e-08	3.3872e-21	2.3629	-S	2L
bard1m	17	17	4.6689e-06	2.9017e-20	1.2256	-S	G
bard2m	6598	-	-	-	-	-	-3
bard3m	-12.6787	-12.6787	2.8072e-08	2.2215e-18	2.3257	-S	2L
bilevel1	0	0*	1.0308e-15	0	3.2387	-S	2L
bilevel1m	-55	5	9.2210e-07	4.2735e-13	2.7696	-S	G
bilevel2	-6600	-3.7876e-12	9.7733e-06	1.6290e-13	1.0050	-S	G
bilevel2m	-6600	-5.0253e-11	1.1351e-07	5.9365e-18	1.5375	-S	G
bilevel3	-12.6787	-12.6787	9.3752e-05*	3.9834e-17	1.0073	-S	G
bilin	18.4	11.6000	2.7935e-06*	5.8560e-14	0.8956	-S	G
dempe	28.25	-	-	-	-	-	-2
desilva	-1	-1	3.6423e-05*	5.0931e-17	1.1209	-S	G
df1	0	0	1.4835e-15	0	4.0850	-S	2L
ex9.1.1	-13	-13	1.9610e-15	4.0059e-32	3.2145	-S	2L

Tabela 4.2: Continuação

Problema	Melh. Obj.	Obj.	$r(x_{k_f}, y_{k_f})$	Inv.	OCE	Stat.	d
ex9.1.3	-29.2	-29.2*	3.3395e-06	6.9288e-12	1.2872	-S	G
ex9.1.4	-37	-37	2.6066e-14	2.5367e-28	5.0917	-S	2L
ex9.1.5	-1	-1	2.8245e-06	5.8326e-12	1.2975	-S	G
ex9.1.6	-49	-49	2.5609e-08	6.5584e-16	1.9823	-S	2L
ex9.1.7	-26	-6	2.5265e-09	6.3830e-18	1.9663	-S	1L
ex9.1.8	-3.25	-3.25	2.4906e-16	0	3.5947	-S	2L
ex9.1.9	3.1111	3.1111	2.1656e-11	7.2970e-30	$\gg 1$	-S	2L
ex9.1.10	-3.25	-3.25	2.4906e-16	0	3.5947	-S	2L
ex9.2.1	17	17	8.0360e-09	6.4577e-17	2.0427	-S	1L
ex9.2.2	100	99.9978	4.2880e-04*	1.8387e-07	$\gg 1$	-S	1L
ex9.2.3	-55	-	-	-	-	-	-2
ex9.2.4	0.5	4	2.7597e-06	2.5755e-14	1.1537	-S	G
ex9.2.5	6	5*	7.1295e-06	5.0830e-11	12.4017	-S	2L
ex9.2.6	-1	-1	1.6052e-10	2.5767e-20	2.2759	-S	2L
ex9.2.7	17	17	2.3078e-06	5.3258e-12	3.4492	-S	1L
ex9.2.8	1.5	1.5	1.7369e-14	3.0016e-28	3.4953	-S	1L
ex9.2.9	2	2	3.0708e-15	0	4.0582	-S	2L
flp2	0	0*	2.6125e-11	0	3.0439	-S	2L
gnash10	-230.823	-230.8232*	1.7229e-08	2.9684e-16	1.8528	-S	2L
gnash11	-129.912	-129.9119*	3.1018e-07	9.5895e-14	2.2105	-S	2L
gnash12	-36.9331	-36.9331*	1.0020e-08	1.0037e-16	1.9667	-S	2L
gnash13	-7.06178	-7.0618*	7.6253e-10	5.8095e-19	1.9597	-S	2L
gnash14	-0.17906	-0.1790*	1.6594e-09	2.7505e-18	1.8395	-S	2L
gnash15	-354.699	-354.6992*	1.5998e-06	2.4601e-12	1.2190	-S	2L
gnash16	-241.442	-241.4420*	1.8034e-08	3.0628e-16	1.5652	-S	2L
gnash17	-90.7491	-90.7491*	9.7761e-06	1.2912e-11	1.1325	-S	2L
gnash18	-25.6982	-25.6989*	8.9010e-06	7.5462e-11	1.0155	-S	2L
gnash19	-6.11671	-6.1167*	1.0105e-06	1.0208e-12	2.4193	-S	2L
gauvin	20	20	4.1655e-07	6.7554e-14	13.1700	-S	2L
hakonsen	24.3668	-	-	-	-	-	-3
jr1	0.5	0.5	2.1583e-09	9.2963e-19	1.7728	-S	2L
jr2	0.5	0.5	6.2008e-06	6.2760e-12	4.9431	-S	1L
kth1	0	4.6025e-15*	3.9066e-13	0	$\gg 1$	-S	G
kth2	0	0*	1.4824e-16	0	3.5080	-S	2L

Tabela 4.2: Continuação

Problema	Melh. Obj.	Obj.	$r(x_{k_f}, y_{k_f})$	Inv.	OCE	Stat.	d
kth3	0.5	0.5	8.3913e-12	0	113.5966	-S	1L
nash1a	7.89e-30	0	4.0517e-15	1.2622e-29	2.8870	-S	2L
nash1b	7.89e-30	1.9722e-30	2.2648e-14	6.3109e-29	3.6957	-S	2L
nash1c	7.89e-30	-	-	-	-	-	-2
nash1d	7.89e-30	5.4780e-12	4.8219e-06	2.4620e-13	1.1717	-S	1L
nash1e	7.89e-30	3.9443e-31	9.9116e-15	1.2622e-29	3.4278	-S	2L
outrata31	3.2077	3.2077	1.6289e-07	2.6276e-14	2.8044	-S	2L
outrata32	3.4494	3.4494	6.4701e-07	1.8940e-14	2.9587	-S	2L
outrata33	4.60425	4.60425	2.3426e-07	4.5332e-14	2.1572	-S	1L
outrata34	6.59268	6.59268	7.1984e-06	2.0751e-13	3.0285	-S	2L
ralph2	0	6.7034e-29	2.5050e-14	0	3.5377	-S	1L
scholtes1	2	2	1.7786e-17	0	4.1792	-S	2L
scholtes2	15	15	3.2704e-08	0	1.9769	-S	2L
scholtes3	0.5	0.5	0	0	$\gg 1$	-S	1L
scholtes4	-3.07e-7	-	-	-	-	-	-1
scholtes5	1	1	5.7703e-09	0	$\gg 1$	-S	1L
scale1	1	1	0	0	$\gg 1$	-S	1L
scale2	1	1	0	0	$\gg 1$	-S	1L
scale3	1	0.9981	9.7671e-06	9.4937e-11	1.0370	-S	2L
scale4	1	-	-	-	-	-	-1
scale5	100	100	0	0	$\gg 1$	-S	1L
sl1	0.0001	-	-	-	-	-	-3
stackelberg1	-3266.67	-3.2667	4.6710e-07	2.1818e-13	5.0588	-S	2L

Descontando os 8 casos em que o Algoritmo 5 falhou, encontrou uma solução melhor que a solução documentada conhecida em 2 casos, enquanto convergiu para um ponto com maior valor de função objetivo em 4 instâncias. É interessante que, embora a verificação de viabilidade exija apenas a ordem menor de $\tau_s = 1e^{-5}$, na maioria dos casos, obtivemos uma precisão muito maior.

Com respeito à ordem de convergência estimada OCE, temos os seguintes resultados, onde se considera o valor de OCE em relação à última direção de descida d (global ou local).

Tabela 4.3: Relação entre OCE e à última direção de descida d .

	Último global	Último local	Dois últimos local	Total
$1.25 < OCE$	5	15	34	54
$1.1 < OCE \leq 1.25$	3	1	2	6
$OCE \leq 1.1$	3	0	2	5

Os resultados indicam que a última direção d é uma direção de descida global em 11 casos. No entanto, 8 desses casos ainda convergem em uma taxa superlinear. Por comparação, dos 54 problemas para os quais a última direção d é uma direção de descida local, a convergência superlinear ocorre em 52 casos.

Conclusões

Neste trabalho pesquisamos sobre a convergência em problemas MPEC de um método de programação quadrática sequencial estabilizado de segunda ordem, o qual está desenhado para problemas não lineares usuais. A ideia surgiu dos trabalhos feitos com o mesmo objetivo em [40] e [1] usando o método de Lagrangiano Aumentado de primeira e segunda ordem respectivamente.

Inicialmente, fizemos uma revisão dos conceitos e propriedades associados aos problemas MPEC relacionados às condições de qualificação e aos conceitos de estacionariedade. Ressaltamos que na procura de ter uma teoria fechada destes conceitos detectamos um erro numa demonstração feita em [23], onde se caracteriza condições suficientes para que um ponto viável de um problema MPEC satisfaça a condição de qualificação de Abadie e apresentamos um exemplo no qual esta afirmação não é válida.

Num segundo momento, escolhemos um algoritmo baseado no método de programação quadrática sequencial estabilizado de segunda ordem desenvolvido em [30]. A análise de convergência deste algoritmo, feita em [30] e [31], se baseia em considerar que o número de iterações de tipo O e V geradas pelo algoritmo é infinito. Porém, para garantir a convergência a pontos M- e S-estacionários, foi necessário modificá-lo. Mostramos que com a referida modificação a hipótese de considerar o número de iterações de tipo O e V infinito também se justificava, fato que não alterou as propriedades de convergência do algoritmo.

A análise da convergência foi feita sobre pontos de acumulação viáveis para o Problema (4.1) de sequências geradas pelo Algoritmo 5, considerando que o conjunto de iterações de tipo O e V é infinito. Uma primeira análise dos pontos AKKT gerados pelo algoritmo, permitiu fixar uma sequência de multiplicadores aproximados $\{\hat{y}_k\}$ associada aos pontos de acumulação. No Algoritmo 5, não existe uma fórmula explícita para calcular os multiplicadores, pois os iterandos aparecem como soluções de um problema quadrático no qual a solução é um vetor formado por variáveis primais e duais. Nesse sentido, foi necessário encontrar uma fórmula explícita para a atualização dos multiplicadores de Lagrange, fato que junto à sequência $\{\hat{y}_k\}$ nos permitiu concluir a resultados importantes nesta tese.

O primeiro está baseado em [1] e mostra convergência a pontos S-estacionários considerando a condição MPEC-RCPLD e uma propriedade de limitação vinculada ao multiplicador da restrição de complementaridade; este resultado é equivalente ao feito em

[1] para Lagrangiano Aumentado. O segundo está baseado em [40] e mostra convergência a pontos M-estacionários considerando a condição MPEC-LICQ, sendo equivalente ao feito em [1] para Lagrangiano Aumentado. No final discutimos o comportamento do algoritmo proposto em problemas da coleção MacMPEC para validar nossos resultados, onde se ressalta que na maioria dos problemas analisados foi encontrada uma solução igual à solução conhecida, além de ter uma taxa de convergência superlinear.

Como sugestão para trabalhos futuros podemos citar o estudo de uma teoria de “*error bound*” em problemas MPEC, considerando a teoria de multifunções para equações generalizadas, onde os problemas MPEC são casos particulares. Uma possível função proximal Lipschitz $\Phi : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ associada a pontos M-estacionários com a propriedade

$$\Phi(a, b) = 0 \iff ab = 0 \text{ ou } (a \geq 0 \text{ e } b \geq 0)$$

pode ser definida por $\Phi(a, b) = g(g(a, b), h(a, b))$ com $g(a, b) = \|(a, b)\| - |a + b|$ e $h(a, b) = |\min(a, 0)| + |\min(b, 0)|$. Esta teoria permitiria estabelecer um algoritmo baseado em Φ para encontrar pontos que são aproximadamente M-estacionários.

Em relação ao erro encontrado em [23], fica em aberto como caracterizar condições necessárias e suficientes para garantir que um ponto viável de um problema MPEC satisfaça a condição de qualificação de Abadie. Além disso esse resultado poderia ser adaptado para uma análise similar considerando a condição de qualificação de Guignard.

Uma outra possibilidade é que a teoria e o algoritmo desenvolvidos poderiam ser aplicados a problemas multiobjetivos com restrições de equilíbrio. Sob certas hipóteses, é possível converter um problema multiobjetivo em um problema mono objetivo [49].

Referências

- [1] R. Andreani, L.D. Secchin and P.J.S. Silva. Convergence properties of a second order augmented Lagrangian method for mathematical programs with complementarity constraints. *SIAM J. Optim.*, 28(3): 2574-2600, 2018.
- [2] R. Andreani, E.G. Birgin, J.M. Martínez and M.L. Schuverdt. Second-order negative-curvature methods for box-constrained and general constrained optimization. *Comput Optim Appl* 45(2): 209-236, 2010.
- [3] R. Andreani, E.G. Birgin, J.M. Martínez and M.L. Schuverdt. On augmented Lagrangian methods with general lower-level constraints. *SIAM J. Optim.*, 18(4):1286-1309, 2007.
- [4] R. Andreani, P.J.S. Silva, G. Haeser and A. Ramos. A second-order sequential optimality condition associated to the convergence of optimization algorithms. *IMA Journal of Numerical Analysis* 37(4): 1902-1929, 2017.
- [5] R. Andreani, G. Haeser, M. L. Schuverdt, and P. J. S. Silva. Two new weak constraint qualifications and applications. *SIAM J. Optim.*, 22(3):1109-1135, 2012.
- [6] R. Andreani, J. M. Martínez, and M. L. Schuverdt. On second-order optimality conditions for nonlinear programming. *Optimization*, 56(5-6):529-542, 2007.
- [7] R. Andreani, G. Haeser and J. M. Martínez. On sequential optimality conditions for smooth constrained optimization. *Optimization*, 60(5), 627-641, 2011.
- [8] R. Andreani, G. Haeser, M. L. Schuverdt, and P. J. S. Silva. A relaxed constant positive linear dependence constraint qualification and applications. *Math. Program.*, 135(1):255-273, 2012.
- [9] R. Andreani, G. Haeser, L.D. Secchin and P.J.S. Silva. New sequential optimality conditions for mathematical problems with complementarity constraints and algorithmic consequences, *Optimization Online*, 2018.
- [10] R. Andreani, C. Dunder, and J. M. Martínez. Nonlinear-programming reformulation of the order-value optimization problem. *Math. Meth Oper. Res.*, 61(3):365-384, 2005.

- [11] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty. *Nonlinear Programming Theory and Algorithms*, John Wiley, New York, 3rd edition, 2006.
- [12] E. Birgin and J. Martínez. *Practical Augmented Lagrangian Methods for Constrained Optimization*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2014.
- [13] R. W. Cottle, J. S. Pang, and R. E. Stone. *The Linear Complementarity Problem. Computer Science and Scientific Computing*, Academic Press, San Diego, 1992.
- [14] Y. Chen and M. Florian. The nonlinear bilevel programming problem: Formulations, regularity and optimality conditions. *Optimization*, 32(3):193-209, 1995.
- [15] F. E. Curtis and J. Nocedal. Flexible penalty functions for nonlinear constrained optimization. *IMA Journal of Numerical Analysis*, 28(4):749-769, 2008.
- [16] S. Dempe. *Foundations of Bilevel Programming*, volume 61 of *Nonconvex Optimization and Its Applications*. Springer US, 1 edition, 2002.
- [17] S. Dempe. Annotated bibliography on bilevel programming and mathematical programs with equilibrium constraints. *Optimization*, 52(3):333-359, 2003.
- [18] F. Facchinei, H. Jiang, and L. Qi. A smoothing method for mathematical programs with equilibrium constraints. *Math. Program.*, 85(1):107-134, 1999.
- [19] M. C. Ferris and J. S. Pang. Engineering and economic applications of complementarity problems. *SIAM Rev.*, 39(4):669-713, 1997.
- [20] A. Fischer. Local behavior of an iterative framework for generalized equations with nonisolated solutions. *Math. Program. Ser. A*, 94(1):91-124, 2002.
- [21] M.L. Flegel and C. Kanzow. Abadie-type constraint qualification for mathematical programs with equilibrium constraints. *J. Optim. Theory Appl*, 124(3):595-614, 2005.
- [22] M. L. Flegel and C. Kanzow. A direct proof for M-stationarity under MPEC-GCQ for mathematical programs with equilibrium constraints, pp 111-122. In: Dempe S., Kalashnikov V. (eds) *Optimization with Multivalued Mappings. Springer Optimization and Its Applications*, vol 2. Springer, Boston, MA.2006.
- [23] M. L. Flegel and C. Kanzow. On the Guignard constraint qualification for mathematical programs with equilibrium constraints. *Optimization*, 54(6):517-534, 2005.
- [24] M. L. Flegel and C. Kanzow. A Fritz-John approach to first order optimality conditions for mathematical programs with equilibrium constraints, *Optimization*, 52(3): 277-286, 2003.

- [25] M. L. Flegel and C. Kanzow. On M-stationarity for mathematical programs with equilibrium constraints, *Journal of Mathematical Analysis and Applications*, 310:286-302.,2005.
- [26] R. Fletcher, S. Leyffer, D. Ralph, and S. Scholtes. Local convergence of SQP methods for mathematical programs with equilibrium constraints. *SIAM J. Optim.*, 17(1):259-286, 2006.
- [27] M. Fukushima and G. H. Lin. Smoothing methods for mathematical programs with equilibrium constraints, in Proceedings of the *International Conference on Informatics Research for Development of Knowledge Society Infrastructure (ICKS'04)*, IEEE Computer Society, 206-213, 2004.
- [28] C. Geiger and C. Kanzow, *Theorie und Numerik restringierter Optimierungsaufgaben*, Springer-Verlag, Berlin/Heidelberg, 2002.
- [29] P. E. Gill and D. P. Robinson. A globally convergent stabilized SQP method. *SIAM J. Optim.*, 23(4):1983-2010, 2013.
- [30] P. E. Gill, V. Kungurtsev, and D. P. Robinson. A stabilized SQP method: Global convergence. *IMA Journal of Numerical Analysis*, 37(1):407-443, 2016.
- [31] P. E. Gill, V. Kungurtsev, and D. P. Robinson. A stabilized SQP method: Superlinear convergence. Ser. A, *Math. Program.*,163(1-2):369-410, 2017.
- [32] L. Guo, G. Lin, J. J. Ye. Stability analysis for parametric mathematical programs with geometric constraints and its applications, *SIAM J. Optim.*, 22(3):1151-1176, 2012.
- [33] L. Guo, G. H. Lin, D. Zhang, and D. Zhu. An MPEC reformulation of an EPEC model for electricity markets. *Oper. Res. Lett.*, 43(3):262-267, 2015.
- [34] L. Guo, G. H. Lin, and J. J. Ye. Second-order optimality conditions for mathematical programs with equilibrium constraints. *J. Optim. Theory Appl.*, 158(1):33-64, 2013.
- [35] L. Guo and G. H. Lin. Notes on some constraint qualifications for mathematical programs with equilibrium constraints. *J. Optim. Theory Appl.*, 156(3):600-616, 2013.
- [36] R. Henrion, A. Jourani, and J. Outrata. On the calmness of a class of multifunctions. *SIAM J. Optim.*, 13(2):603-618, 2002.
- [37] T. Hoheisel, C. Kanzow, and A. Schwartz. Convergence of a local regularization approach for mathematical programs with complementarity or vanishing constraints. *Optimization Methods and Software*, 27(3):483-512, 2012.

- [38] T. Hoheisel, C. Kanzow, and A. Schwartz. Theoretical and numerical comparison of relaxation methods for mathematical programs with complementarity constraints. *Math. Program.*, 137(1-2):257-288, 2013.
- [39] X. M. Hu and D. Ralph. Convergence of a penalty method for mathematical programming with complementarity constraints. *J. Optim. Theory Appl.*, 123(2):365-390, 2004.
- [40] A. F. Izmailov, M. V. Solodov and E. I. Uskov. Global convergence of augmented Lagrangian methods applied to optimization problems with degenerate constraints, including problems with complementarity constraints. *SIAM J. Optim.*, 22(4):1579-1606, 2012.
- [41] A. F. Izmailov and M. V. Solodov, *Newton-Type Methods for Optimization and Variational Problems*, Springer, New York, 2014.
- [42] A. F. Izmailov and M. V. Solodov. *Otimização: Métodos Computacionais*, volume 2. IMPA, Rio de Janeiro, 2007.
- [43] A. F. Izmailov. Solution sensitivity for Karush-Kuhn-Tucker systems with non-unique Lagrange multipliers. *Optimization*, 59(5):747-775, 2010.
- [44] A. Kadrani, J. P. Dussault, and A. Benchakroun. A globally convergent algorithm for MPCC. *EURO J. Comput. Optim.*, 3(3):263-296, 2015.
- [45] A. Kadrani, J. P. Dussault, and A. Benchakroun. A new regularization scheme for mathematical programs with complementarity constraints. *SIAM J. Optim.*, 20(1):78-103, 2009.
- [46] C. Kanzow and A. Schwartz. Mathematical programs with equilibrium constraints: Enhanced Fritz John-conditions, new constraint qualifications, and improved exact penalty results. *SIAM J. Optim.*, 20(5):2730-2753, 2010.
- [47] C. Kanzow and A. Schwartz. A new regularization method for mathematical programs with complementarity constraints with strong convergence properties. *SIAM J. Optim.*, 23(2):770-798, 2013.
- [48] S. Leyffer, G. López, and J. Nocedal. Interior methods for mathematical programs with complementarity constraints. *SIAM J. Optim.*, 17(1):52-77, 2009.
- [49] S. Leyffer. A complementarity constraint formulation of convex multiobjective optimization problems. *INFORMS Journal on Computing*, 21(2):257-267, 2009.
- [50] G. H. Lin and M. Fukushima. A modified relaxation scheme for mathematical programs with complementarity constraints. *Ann. Oper. Res.*, 133(1-4):63-84, 2005.

- [51] G. H. Lin, L. Guo, and J. J. Ye. Solving mathematical programs with equilibrium constraints. *J. Optim. Theory Appl.* 166(1):234-256, 2015
- [52] X. Liu and J. Sun. Generalized stationary points and an interior-point method for mathematical programs with equilibrium constraints. *Math. Program., Ser. A*, 101(1):231-261, 2004.
- [53] Q. Luo, J. S. Pang, and D. Ralph. Mathematical Programs with Equilibrium Constraints, Cambridge University Press, Cambridge, UK, 1996.
- [54] B. S. Mordukhovich. Generalized differential calculus for nonsmooth and set-valued mappings, *Journal of Mathematical Analysis and Applications*, 183(1):250-288, 1994.
- [55] J. V. Outrata, M. Kovara, and J. Zowe. Nonsmooth Approach to Optimization Problems with Equilibrium Constraints, *Nonconvex Optimization and its Applications*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998.
- [56] J. V. Outrata. Optimality conditions for a class of mathematical programs with equilibrium constraints, *Mathematics of Operations Research*, 24(3):627-644, 1999.
- [57] J. V. Outrata, A generalized mathematical program with equilibrium constraints, *SIAM J. Control Optim.*, 38(5):1623-1638, 2000.
- [58] J.-S. Pang and M. Fukushima. Complementarity constraint qualifications and simplified B-stationarity conditions for mathematical programs with equilibrium constraints, *Computational Optimization and Applications* 13(1-3):111-136, 1999.
- [59] D. W. Peterson. A review of constraint qualifications in finite-dimensional spaces, *SIAM Rev.*, 15(3):639-654, 1973.
- [60] A. A. Ribeiro, E. W. Karas. *Otimização Contínua, Aspectos teóricos e computacionais*, Cengage Learning, São Paulo-SP, Brazil, 2013.
- [61] A. Ramos. Mathematical programmes with equilibrium constraints: A sequential optimality condition, new constraint qualifications and algorithmic consequences, *Optimization Online*, 2016.
- [62] A. Ramos. *Tópicos em Condições de Otimalidade para Otimização não Linear. Ph.D. Thesis*, IME-USP, Departamento de Matemática Aplicada, São Paulo-SP, Brazil, 2016.
- [63] D. P. Robinson. Primal-Dual Methods for Nonlinear Optimization. PhD thesis, Department of Mathematics, University of California, San Diego, September 2007.
- [64] R. T. Rockafellar and R. J-B. Wets. Variational Analysis, vol. 317 of *A Series of Comprehensive Studies in Mathematics*, Springer, Berlin, Heidelberg, 1998.

- [65] H. Scheel and S. Scholtes. Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity. *Mathematics of Operations Research*, 25(1):1-22, 2000.
- [66] S. Scholtes. Convergence properties of a regularization scheme for mathematical programs with complementarity constraints. *SIAM J. Optim.*, 11(4):918-936, 2001.
- [67] S. Steffensen and M. Ulbrich. A new relaxation scheme for mathematical programs with equilibrium constraints. *SIAM J. Optim.*, 20(5):2504-2539, 2010.
- [68] D. Stewart and M. Anitescu. Computing optimal controls for friction problems. Talk at the First MPS International Conference on Continuous Optimization, Troy, NY, August 2004.
- [69] S. J. Wright. An algorithm for degenerate nonlinear programming with rapid local convergence. *SIAM J. Optim.*, 15(3):673-696, 2005.
- [70] H. Xu, J. S. Pang, F. Ordóñez, and M. Dessouky. Complementarity models for traffic equilibrium with ridesharing. *Transportation Research Part B: Methodological*, 81(1):161-182, 2015.
- [71] X. Q. Yang and X. X. Huang. Lower-order penalty methods for mathematical programs with complementarity constraints. *Optimization Methods and Software*, 19(6):693-720, 2004.
- [72] X.Q. Yang and X.X. Huang. Convergence analysis of an augmented Lagrangian method for mathematical programs with complementarity constraints. *Nonlinear Anal.*, 63(5-7):e2247- e2256, 2005.
- [73] J. J. Ye, and D. L. Zhu. Optimality conditions for bilevel programming problems. *Optimization*, 33(1): 9-27, 1995
- [74] J. J. Ye Necessary and sufficient optimality conditions for mathematical programs with equilibrium constraints, *Journal on Mathematical Analysis and Applications*, 307(1):350-369, 2005.
- [75] J. J. Ye, D. L. Zhu, and Q. J. Zhu. Exact penalization and necessary optimality conditions for generalized bilevel programming problems, *SIAM J. Optim.*, 7(2):481-507, 1997.
- [76] J. J. Ye. Constraint qualifications and necessary optimality conditions for optimization problems with variational inequality constraints, *SIAM J. Optim.*, 10(4):943-962, 2000.

- [77] J. J. Ye and X. Y. Ye, Necessary optimality conditions for optimization problems with variational inequality constraints, *Mathematics of Operations Research*, 22(4):977-997, 1997.

Apêndice (Algoritmos)

Algoritmo 1 (estima uma direção de curvatura negativa $s_k^{(1)}$)

```

1: Dados  $x_k, y_k, \mu_{k-1}^R, J_k, H_k$ 
2:   Computar  $H_{\mathcal{F}_\epsilon}$  e  $J_{\mathcal{F}_\epsilon}$  como submatrizes de  $H_k$  e  $J_k$  associadas a  $\mathcal{F}_\epsilon(x_k, y_k, \mu_{k-1}^R)$ ;
3:   se  $H_{\mathcal{F}_\epsilon} + \frac{1}{\mu_{k-1}^R} J_{\mathcal{F}_\epsilon}^T J_{\mathcal{F}_\epsilon}$  é positiva semidefinida então
4:      $\xi_k = 0, u_k^{(1)} = 0$  e  $w_k^{(1)} = 0$ ;
5:   caso contrário
6:     Computar  $u_{\mathcal{F}_\epsilon} \neq 0$  que cumpra (3.19);
7:      $\xi_k = -u_{\mathcal{F}_\epsilon}^T (H_{\mathcal{F}_\epsilon} + (1/\mu_{k-1}^R) J_{\mathcal{F}_\epsilon}^T J_{\mathcal{F}_\epsilon}) u_{\mathcal{F}_\epsilon} / \|u_{\mathcal{F}_\epsilon}\|^2 > 0$ ;
8:      $u_k^{(1)} = 0$ ;  $[u_k^{(1)}]_{\mathcal{F}_\epsilon} = u_{\mathcal{F}_\epsilon}$ ; e
9:      $w_k^{(1)} = -(1/\mu_{k-1}^R) J_k u_k^{(1)}$ ;
10:  fim
11:   $s_k^{(1)} = (u_k^{(1)}, w_k^{(1)})$ ;
12:  saída  $(s_k^{(1)}, \xi_k)$ ;
13: fim

```

Algoritmo 2 (encontra uma direção de descida d_k)

```

1: Dados  $x_k, y_k, \mu_k^R, y_k^E, J_k, H_k$ 
2:   Constantes:  $0 < \lambda < \min \{\bar{\gamma}, 1 - \bar{\gamma}\} < 1$ ;
3:    $B = B(x_k, y_k; \mu_k^R)$ ; Computar uma matriz definida positiva  $\hat{B}$  (3.33) a partir de  $B$ ;
4:    $\nabla M_k = \nabla M(x_k, y_k; y_k^E, \mu_k^R)$ ;  $t_k = r(x_k, y_k)^\lambda$ ;  $v_k = (x_k, y_k)$ ;
5:    $[\hat{v}_k^{(0)}]_{\mathcal{A}_\epsilon} = 0$ ;  $[\hat{v}_k^{(0)}]_{\mathcal{F}_\epsilon} = [v_k]_{\mathcal{F}_\epsilon}$ ;
6:   se ( $B_{\mathcal{F}_\epsilon}$  é definida positiva e  $v_k$  é uma V-O iteração) então
7:      $[\Delta \hat{v}_k^{(0)}]_{\mathcal{A}_\epsilon} = 0$ ;  $B_{\mathcal{F}_\epsilon} [\Delta \hat{v}_k^{(0)}]_{\mathcal{F}_\epsilon} = -[\nabla Q_k(\hat{v}_k^{(0)})]_{\mathcal{F}_\epsilon}$ ;  $\hat{v}_k = \hat{v}_k^{(0)} + \Delta \hat{v}_k^{(0)}$ ;
8:      $d_k = \hat{v}_k - v_k$ ;
9:     se ( $v_k + d_k$  é viável e  $\nabla M_k^T d_k < 0$  e  $[\nabla Q_k(v_k + d_k)]_{\mathcal{A}_\epsilon} \geq -t_k e$ ) então
10:       saída  $d_k$ ; (direção de descida local)
11:   fim
12:   fim
13:    $[\Delta \hat{v}_k^{(0)}]_{\mathcal{A}_\epsilon} = 0$ ; Resolver  $\hat{B}_{\mathcal{F}_\epsilon} [\Delta \hat{v}_k^{(0)}]_{\mathcal{F}_\epsilon} = -[\nabla \hat{Q}_k(\hat{v}_k^{(0)})]_{\mathcal{F}_\epsilon}$ ;
14:   Computar  $\hat{\alpha}_0 \geq 0$  e  $\hat{v}_k^{(1)}$  tais que  $\hat{v}_k^{(1)} = \hat{v}_k^{(0)} + \hat{\alpha}_0 \Delta \hat{v}_k^{(0)}$  é viável;
15:   Resolver o problema convexo (3.32) para  $\hat{v}_k$  considerando  $\hat{v}_k^{(1)}$  como ponto inicial;
16:    $d_k = \hat{v}_k - v_k$ ; (direção de descida global)
17:   saída  $d_k$ ;
18: fim

```

- 1: **Dados** $s_k^{(1)}, \xi_k, x_k, d_k, J_k, H_k$
- 2: $s_k^{(2)} = \begin{cases} -s_k^{(1)}, & \text{se } \nabla M(v_k, y_k^E, \mu_k^R)^T s_k^{(1)} > 0; \\ s_k^{(1)}, & \text{caso contrário;} \end{cases}$
- 3: Considerar p_k e $u_k^{(2)}$ as primeiras n componentes de d_k e $s_k^{(2)}$;
- 4: $\sigma_k = \arg \max_{\sigma \geq 0} \left\{ \sigma : x_k + p_k + \sigma u_k^{(2)} \geq 0, \left\| \sigma u_k^{(2)} \right\| \leq \max(\xi_k, \|p_k\|) \right\}$;
- 5: $s_k = \sigma_k s_k^{(2)}$;
- 6: **saída** s_k ;
- 7: **fim**

Algoritmo 4 (busca linear flexível)

- 1: **Dados** $d_k, s_k, y_k^E, \mu_k, \mu_k^R, \mu_{k-1}^R, \ell_k, J_k, H_k$
- 2: **Constantes:** $\gamma_S \in (0, \frac{1}{2})$;
- 3: Computar $\nabla M = \nabla M(x_k, y_k, y_k^E, \mu_k^R)$;
- 4: **se** $s_k = 0$ **e** $d_k = 0$ **então**
- 5: $\alpha_k = 1$;
- 6: **caso contrário** ($d_k \neq 0$ **ou** $\nabla M^T s_k < 0$ **ou** $\mu_k^R = \mu_{k-1}^R$) **então**
- 7: $\alpha_k = 1$;
- 8: **enquanto** $\rho_k(\alpha_k; \mu_k^R, \ell_k) < \gamma_S$ **e** $\rho_k(\alpha_k; \mu_k, \ell_k) < \gamma_S$
- 9: $\alpha_k = \frac{1}{2}\alpha_k$;
- 10: **fim**
- 11: **caso contrário** [$d_k = 0, s_k \neq 0, \xi_k > 0$]
- 12: $\xi_k^R = -s_k^T \nabla^2 M(x_k, y_k, y_k^E, \mu_k^R) s_k / \|u_k\|^2$; [Por definição, $s_k = (u_k, w_k)$]
- 13: **se** $\xi_k^R > \gamma_S \xi_k$ **então**
- 14: $\alpha_k = 1$;
- 15: **enquanto** $\rho_k(\alpha_k, \mu_k^R, \ell_k) < \gamma_S$ **e** $\rho_k(\alpha_k, \mu_k, \ell_k) < \gamma_S$
- 16: $\alpha_k = \frac{1}{2}\alpha_k$;
- 17: **fim**
- 18: **caso contrário**
- 19: $\alpha_k = 0$;
- 20: **fim**
- 21: **fim**
- 22: **saída** $\alpha_k \geq 0$
- 23: **fim**

Algoritmo 5 (PQS primal-dual de segunda ordem)

- 1: **Dados** x_1, y_1, μ_0^R, μ_1
- 2: **Constantes:** $\{\tau_s, \gamma_s\} \subset (0, \frac{1}{2})$, $0 < \bar{\gamma} < 1$, $y_{\max} > 0$ e $0 < \epsilon_a \ll 1$;
- 3: escolher $y_0^E \in \mathbb{R}^m$, $\{\tau_0, \phi_{V,0}^{\max}, \phi_{O,0}^{\max}\} \subset (0, \infty)$ e $0 < \mu_0^R \leq \mu_1 < \infty$;
- 4: $k = 1$;
- 5: **repita**
- 6: Computar o conjunto ϵ -livre $\mathcal{F}_\epsilon(x_k, y_k, \mu_{k-1}^R)$ de (3.18)
- 7: $J_k = J(x_k)$; $H_k = H(x_k, y_k)$;
- 8: $(s_k^{(1)}, \xi_k) = \text{menor_curvatura}(x_k, y_k, \mu_{k-1}^R, J_k, H_k)$; [Algoritmo 1]
- 9: Calcular $r(x_k, y_k)$ de (3.2);
- 10: **se** (vale condição de parada (3.39)) **então**
- 11: **saída** o ponto KKT de segunda ordem (x_k, y_k) ; (Definição 1.14);
- 12: **fim**
- 13: **se** $(\phi_V(x_k, y_k) \leq \frac{1}{2}\phi_{V,k-1}^{\max})$ e $\|y_k\| \leq y_{\max}$ **então** [V-iteração]
- 14: $\phi_{V,k}^{\max} = \frac{1}{2}\phi_{V,k-1}^{\max}$; $y_k^E = y_k$; $\tau_k = \tau_{k-1}$;
- 15: Atualizar μ_k^R como em (3.26);
- 16: **se em outro caso** $(\phi_O(x_k, y_k) \leq \frac{1}{2}\phi_{O,k-1}^{\max})$ e $\|y_k\| \leq y_{\max}$ [O-iteração]
- 17: $\phi_{O,k}^{\max} = \frac{1}{2}\phi_{O,k-1}^{\max}$; $y_k^E = y_k$; $\tau_k = \tau_{k-1}$;
- 18: Atualizar μ_k^R como em (3.26);
- 19: **se em outro caso** $((x_k, y_k)$ satisfaz (3.28)) **então** [M-iteração]
- 20: Atualizar y_k^E como em (3.29); $\tau_k = \frac{1}{2}\tau_{k-1}$;
- 21: Atualizar μ_k^R como em (3.30);
- 22: **caso contrário**
- 23: $y_k^E = y_{k-1}^E$; $\tau_k = \tau_{k-1}$; $\mu_k^R = \mu_{k-1}^R$; [F-iteração]
- 24: **fim**
- 25: **se** (vale condição de parada (3.41)) **então**
- 26: **saída** com o ponto estacionário aproximado não viável x_k ;
- 27: **fim**
- 28: Computar o conjunto ϵ -livre $\mathcal{F}_\epsilon(x_k, y_k, \mu_k^R)$ de (3.18);
- 29: $d_k = \text{direção_descida}(x_k, y_k, \mu_k^R, J_k, H_k)$; [Algoritmo 2]
- 30: $s_k = \text{direção_curvatura_negativa}(s_k^{(1)}, \xi_k, x_k, d_k, J_k, H_k)$; [Algoritmo 3]
- 31: **se** $(d_k \neq 0$ e $s_k = 0$ e (x_k, y_k) é uma V-O iteração) **então**
- 32: $\ell_k = 1$;
- 33: **caso contrário**
- 34: $\ell_k = 2$;
- 35: **fim**
- 36: $\mu_k = \max(\mu_k^R, \mu_k)$;
- 37: $\alpha_k = \text{busca_linear_flexível}(d_k, s_k, y_k^E, \mu_k, \mu_k^R, \mu_{k-1}^R, \ell_k, J_k, H_k)$; [Algoritmo 4]
- 38: atualizar μ_{k+1} como em (3.38);
- 39: $v_{k+1} = (x_{k+1}, y_{k+1}) = v_k + \alpha_k d_k + \alpha_k s_k$;

```
40:          $k = k + 1$ ;  
41:     fm  
42: fm
```